Proceedings of the 1st Virtual Control Conference VCC 2010

Stoustrup, Jakob; Leth, John-Josef; Schiøler, Henrik

Publication date:
2010

Document Version
Også kaldet Forlagets PDF

VCC-10

AALBORG UNIVERSITY PRESS

Proceedings

of the

1st Virtual Control Conference

VCC 2010

Published by Aalborg University Press, 2010
http://forlag.hum.aau.dk/
ISBN: 978-87-7307-999-7

Organizing Committee

Chair:
Henrik Schiøler,
Vice-chair:
Jakob Stoustrup,
Members:
Jens Dalsgaard Nielsen,
Kirsten Mølgaard Nielsen,
John Leth,
Jan Bendtsen,
Per P. Madsen.

International Program Committee

Frank Allgöwer,
Karolos Grigoriadis,
Tarunraj Singh,
Tariq Samad,
Ian R. Petersen,
Kristin Pettersen,
Claudio De Persis,
Kevin L. Moore,
N. Eva Wu,
Faryar Jabbari,
Anders Rantzer,
Karl H. Johansson,
Kemin Zhou.

Contents

1 Reference Tracking and Profit Optimization of a Power Plant

Reference Tracking and Profit Optimization of a Power Plant

Martin Kragelund, John Leth, and Rafał Wisniewski

***Abstract*— In this paper we discuss two different methods for implementing reference tracking in a profit optimization problem of a power plant. It is shown that tracking included as a side constraint results in an significant tracking error only when the reference gradient is large. When tracking is included in the cost function, as a quadratic term, the reference is tracked with a small accumulated error. Finally, the two methods are compared both in terms of tracking performance and computational burden.**

I. Introduction

Traditional thermal power plants, i.e., coal, gas, or oil fired power plants, have been studied in details [1]. In brief, a thermal power plant functions by burning a fuel in a boiler which evaporates water to steam under high pressure. The stream then drives a turbine generating electrical power which is delivered to the electrical grid.

A thermal power plant is modeled by first principle in [2], where the considered fuel is coal dust which is fed by four coal mills grinding the raw coal. The detailed model in [2] was used to establish an observer for the flow of coal into the boiler to improve the control of the coal mills. Simpler models for system control are presented in [3], where the different methods for changing the output from the complete portfolio of DONG Energy in Denmark are described. The means of changing the output is denoted an effectuator in [3], and the models of typical effectuators in a power plant are derived. An example of an effectuator is the boiler load in a thermal power plant which can be modeled as a 3rd order system.

In production economics the possible outputs from a production unit or "firm" are identified and called the production set [4, Chapter 5]. The production units are seen as black boxes which are capable of transforming some goods (input) to other goods (output). Some assumptions are often made about the production set e.g. *No free lunch* and *Free disposal*, i.e. the production set, Y, cannot contain $\mathbb{R}^l_+$ as this would yield production of some quantity without consumption and the company can absorb any additional input without reducing the output. In [4, Chapter 5] it is concluded that the objective of a company is to maximize its profit, which at first seems reasonable. However, it is possible to imagine companies which have the objective of maximizing sales revenue or the size of the company, but if the company is owned by the consumers in a market they will agree that profit maximization is preferable regardless of their own preference function.

The electricity market place for Nordic Countries is called Nord Pool. Here the price of electricity, as known by the average electricity consumer, is negotiated. Furthermore Nord Pool regulates related to the quality of the power deliverance. These quantities are traded on the hourly spot market, elspot. The transmission system operator maintains the energy balance. In other words it takes care of the situation when a power plant delivers too much or too little electricity to the grid than agreed. To ensure that sufficient reserve capacity is available the transmission system operator pays two prices, an up price and a down price, i.e. price for producing more or less electricity than previously agreed.

The data from Nord Pool has been used before to schedule the usage of hydro power plant in Norway such that the production plan commitment of the current day is fulfilled while maximizing the profit of the hydro plant [6].

This work focus on two different methods for including reference tracking into the design of an optimal profit strategy for a power plant, using coal, gas and oil, under the consideration of two business objectives, efficiency and controllability.

A. Outline

In Section II the plant dynamics, business objectives and profit function are described. In Section III the continuous optimization problem is formulated without reference tracking. For this purpose a discrete formulation of the problem is derived. In Section IV and Section V the reference tracking is included into the optimization problem as a side constraint and as a quadratic term in cost function, respectively. Section VI contain a discussion of the methods for implementing the tracking.

II. Plant Model

In this section a model of the power plant considered in this work is presented. The plant is capable of using three different fuel systems; coal, gas, and oil. For further details about the presented models and quantities the reader is referred to [7]–[10].

This work is supported by The Danish Research Council for Technology and Production Sciences.

The second author is financed by The Danish Council for Technology and Innovation.

The authors are all with the Department of Electronic Systems, Aalborg University, Fredrik Bajers Vej 7C, 9220 Aalborg Ø, Denmark {mkr, jjl, raf}@es.aau.dk

A. Plant Dynamics

The fuel flow, $\boldsymbol{x}(t)$ $[kg/s]$, into the power plant is governed by third order differential equations (these equations also include the power plant dynamics). The control signal to the valves controlling these flows is denoted $\boldsymbol{u} = (u_c, u_g, u_o) \in U,\ U = \left\{ \boldsymbol{v} \in \mathbb{R}^3_+ \mid 0 \leq \boldsymbol{v}^T \boldsymbol{e}_u \leq 400 \right\}$, where $\boldsymbol{e}_u = (10.77, 18.87, 15.77)$ $[kg/s]$, and the dynamics is given by

$$\begin{aligned} \dot{\boldsymbol{z}}(t) &= \boldsymbol{A}\boldsymbol{z}(t) + \boldsymbol{B}\boldsymbol{u}(t) \\ \boldsymbol{x}(t) &= \boldsymbol{C}\boldsymbol{z}(t), \end{aligned} \tag{1}$$

where

$$\boldsymbol{A} = \begin{bmatrix} \boldsymbol{A}_c & \boldsymbol{0}_{3x3} & \boldsymbol{0}_{3x3} \\ \boldsymbol{0}_{3x3} & \boldsymbol{A}_g & \boldsymbol{0}_{3x3} \\ \boldsymbol{0}_{3x3} & \boldsymbol{0}_{3x3} & \boldsymbol{A}_o \end{bmatrix}, \quad \boldsymbol{A}_i = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ h_{i_1} & h_{i_2} & h_{i_3} \end{bmatrix},$$
$$\boldsymbol{B} = \begin{bmatrix} \boldsymbol{B}_c & \boldsymbol{0}_{3x1} & \boldsymbol{0}_{3x1} \\ \boldsymbol{0}_{3x1} & \boldsymbol{B}_g & \boldsymbol{0}_{3x1} \\ \boldsymbol{0}_{3x1} & \boldsymbol{0}_{3x1} & \boldsymbol{B}_o \end{bmatrix}, \quad \boldsymbol{B}_i = \begin{bmatrix} 0 \\ 0 \\ h_{i_0} \end{bmatrix},$$
$$\boldsymbol{C} = \begin{bmatrix} \boldsymbol{C}_1 & \boldsymbol{0}_{1x3} & \boldsymbol{0}_{1x3} \\ \boldsymbol{0}_{1x3} & \boldsymbol{C}_1 & \boldsymbol{0}_{1x3} \\ \boldsymbol{0}_{1x3} & \boldsymbol{0}_{1x3} & \boldsymbol{C}_1 \end{bmatrix}, \quad \boldsymbol{C}_1 = \begin{bmatrix} 1 & 0 & 0 \end{bmatrix},$$

and $h_{i_j}, i \in \mathcal{I}$, are constants describing the dynamics of the three fuel systems which are obtained from transfer functions of the form

$$H_i(s) = \frac{1}{(\tau_i s + 1)^3},$$

where $\tau_i, i \in \mathcal{I}$, is 90, 60, and 70, respectively. The three fuel systems may have some shared dynamics but to simplify the model in this work the systems are assumed decoupled.

Functions describing the two business objectives are derived in the following.

B. Efficiency

The efficiency objective, $y_e = y_e(\boldsymbol{z})$, deals with how much electricity is produced from a certain amount of fuel. Three affine functions describing the contribution of the individual fuels to the efficiency objective have been established using measurement data from two Danish power plants and can be expressed as

$$\tilde{\boldsymbol{y}}_e(\boldsymbol{z}) = \tilde{\boldsymbol{Q}}\boldsymbol{z} + \boldsymbol{b}, \tag{2}$$

where

$$\begin{aligned} &\tilde{\boldsymbol{Q}} = \mathrm{diag}(\boldsymbol{e}_x)\boldsymbol{C}, \quad \boldsymbol{e}_x = (10.77, 18.87, 15.77), \\ &\boldsymbol{b} = (-1.76, 1.85, -0.37), \end{aligned}$$

and $\boldsymbol{C}$ defined in (1). The values of $\boldsymbol{e}_x$ and $\boldsymbol{b}$ have been established using measurement data and are measured in $[MJ/kg]$ and $[MW]$ respectively. The energy used for preprocessing the individual fuels is expressed by the b_i's, and the e_{x_i}'s are conversion factors which are a combination of the boiler efficiency and energy storage in the different fuels.

The total amount of efficiency is described by the function

$$Z \to Y_1;\ \boldsymbol{z} \mapsto y_e(\boldsymbol{z}) = \boldsymbol{\gamma}^T \tilde{\boldsymbol{y}}_e(\boldsymbol{z}),$$

where

$$\boldsymbol{\gamma} = (1, 1, 1).$$

C. Controllability

The controllability objective, $y_c = y_c(\boldsymbol{z})$, deals with a measure of how fast the production of electricity can be changed. Allowed changes in the production is limited to a certain gradient depending on the current efficiency. The reason for this limit is a compliance to maximum temperature gradients in the boiler (these have not been explicitly modelled and are therefore indirectly considered by limiting the allowed changes). When using coal it is allowed to change production with 0.133 $[MW/s]$ when running the plant at low and high production and 0.267 $[MW/s]$ in the middle range from 200 $[MW]$ to 360 $[MW]$. When using oil or gas the values are 0.133 $[MW/s]$ and 0.534 $[MW/s]$. If a mixture of the three fuels are used it is assumed that the allowed change is a linear combination of the allowed change of the individual fuels. The controllability objective is, therefore, modelled as

$$Z \to Y_2;\ \boldsymbol{z} \mapsto y_c(\boldsymbol{z}) = \begin{cases} 0.133 & y_e(\boldsymbol{z}) \in S_1 \\ \frac{\boldsymbol{\xi}^T \tilde{\boldsymbol{y}}_e(\boldsymbol{z})}{y_e(\boldsymbol{z})} & y_e(\boldsymbol{z}) \in S_2 \\ 0.133 & y_e(\boldsymbol{z}) \in S_3, \end{cases} \tag{3}$$

where

$$\begin{aligned} &\boldsymbol{\xi} = (0.267, 0.534, 0.534), \quad S_1 = \{s \in \mathbb{R} | 0 \leq s \leq 200\}, \\ &S_2 = \{s \in \mathbb{R} | 200 < s < 360\}, \text{ and} \\ &S_3 = \{s \in \mathbb{R} | 360 \leq s \leq 400\}. \end{aligned}$$

D. Prices

The cost of using the fuel, revenue from production of output, and the profit of operating the power plant can now be determined. The above constructions yields a product (or output) function, y_P, of the system given by

$$y_P : Z \to Y;\ \boldsymbol{z} \mapsto (y_e(\boldsymbol{z}), y_c(\boldsymbol{z})).$$

The growth of cost and growth of revenue for the system are defined by the following functions (both with units in [DKK/s])

$$\begin{aligned} &g_C : Z \to \mathbb{R};\ \boldsymbol{z} \mapsto \boldsymbol{z}^T \boldsymbol{C}^T \boldsymbol{p}_C, \\ &g_R : Y \times \mathbb{R}_+ \to \mathbb{R};\ (\boldsymbol{y}, t) \mapsto \boldsymbol{y}^T \boldsymbol{p}_R(t), \quad \boldsymbol{p}_R(t) > 0, \end{aligned}$$

where $\boldsymbol{p}_C = (1.20, 3.74, 6.00)$ is the price of coal, gas, and oil respectively and

$$\boldsymbol{p}_R(t) = (p_{R1}(t), p_{R2}(t))$$

the price of the efficiency and controllability respectively.[1]

The growth of profit is hence defined by

$$Z \times Y \times \mathbb{R}_+ \to \mathbb{R};\ (\boldsymbol{z}, \boldsymbol{y}, t) \mapsto g_R(\boldsymbol{y}, t) - g_C(\boldsymbol{z}),$$

which for the system yields the function

$$g_P : Z \times \mathbb{R}_+ \to \mathbb{R};\ (\boldsymbol{z}, t) \mapsto g_R(\boldsymbol{y}_P(\boldsymbol{z}), t) - g_C(\boldsymbol{z}).$$

Therefore, the profit is given by

$$P : \mathbb{R}_+ \to \mathbb{R};\ t \mapsto \int_0^t g_P(\boldsymbol{z}(\tau), \tau) d\tau. \tag{4}$$

III. Problem Formulation

Using the above it is now possible to formulated the following optimization problem

$$\begin{array}{ll} \max\limits_{u \in U} & P(T) = \int_0^T g_P(\boldsymbol{z}, t) dt \\ \text{subject to} & \dot{\boldsymbol{z}} = \boldsymbol{A}\boldsymbol{z} + \boldsymbol{B}\boldsymbol{u}, \end{array} \tag{5}$$

and with the additional requirement that, $y_e(\boldsymbol{z}(t))$ should track a predefined reference signal, $y_r(t)$.

For computational reasons the optimization problem above will be simplified by introduction two approximations. One which assumes good reference tracking and one which deals with condition for discretization of (5).

The growth of profit function, g_P, can when $y_e \approx y_r$ be approximated by

$$g_p(\boldsymbol{z}, t) = \boldsymbol{\Theta}(t)\boldsymbol{z} + \tilde{\varphi}(t), \tag{6}$$

where

$$\begin{aligned} \boldsymbol{\Theta}(t) &= p_{R1}(t)\boldsymbol{\gamma}^T\boldsymbol{Q} - p_{R2}(t)\boldsymbol{p}_C^T\boldsymbol{C} + \boldsymbol{\vartheta}(t), \\ \tilde{\varphi}(t) &= p_{R1}(t)\boldsymbol{\gamma}^T\boldsymbol{b} + p_{R2}(t)\zeta(t), \end{aligned}$$

and $\boldsymbol{\vartheta}(t)$ and $\zeta(t)$ makes up for the switching function in the original formulation of the controllability, i.e.,

$$\boldsymbol{\vartheta}(t) = \begin{cases} 0 & y_r(t) \in S_1 \\ \frac{\boldsymbol{\xi}^T\boldsymbol{Q}}{y_r(t)} & y_r(t) \in S_2 \\ 0 & y_r(t) \in S_3, \end{cases}$$

$$\zeta(t) = \begin{cases} 0.133 & y_r(t) \in S_1 \\ \frac{\boldsymbol{\xi}^T\boldsymbol{b}}{y_r(t)} & y_r(t) \in S_2 \\ 0.133 & y_r(t) \in S_3, \end{cases}$$

Hence the assumption $y_e \approx y_r$ enables us to consider the growth of profit (4) as a affine function of the state as is (6). Note that the assumption also implies that the switching condition $y_r(t) \in S_i$ in the expression for $\boldsymbol{\vartheta}$ and ξ are time dependent, this switching condition would be state dependent otherwise.

The time period T is divided into N equally sized time units, h, i.e., $T = Nh$. It is assumed that $\boldsymbol{\Theta}(t)$, $\varphi(t)$, $\boldsymbol{\psi}(t)$, $y_r(t)$ can be approximated by piecewise constant functions for each time step, i.e.,

$$\begin{aligned} \boldsymbol{\Theta}(t) &= \boldsymbol{\Theta}_k, \quad kh < t < (k+1)h, \\ \tilde{\varphi}(t) &= \tilde{\varphi}_k, \quad kh < t < (k+1)h, \\ y_r(t) &= y_{r_k}, \quad kh < t < (k+1)h. \end{aligned}$$

Furthermore, the control will be assumed piecewise constant as customary when digital to analogue conversion is performed using sample-hold circuits.

Using a fact from [11] the continuous time state $\boldsymbol{z}(t)$ in the dynamical system in (5) can be described by

$$\begin{aligned} \boldsymbol{z}(t) &= e^{\boldsymbol{A}t}\boldsymbol{z}_0 + \int_0^t e^{\boldsymbol{A}(t-s)}\boldsymbol{B}\boldsymbol{u}_0(s)ds \\ &= \begin{bmatrix} \boldsymbol{I} & \boldsymbol{0} \end{bmatrix} \exp\left\{ \begin{bmatrix} \boldsymbol{A} & \boldsymbol{B} \\ \boldsymbol{0} & \boldsymbol{0} \end{bmatrix} t \right\} \begin{bmatrix} \boldsymbol{z}_0 \\ \boldsymbol{u}_0 \end{bmatrix}, \end{aligned} \tag{7}$$

where $\boldsymbol{I}$ is an identity matrix with appropriate dimension. Using (7) it is possible to derive the following formula which is used during the discretization of the cost and constraint

$$\begin{aligned} \int_0^h e^{\boldsymbol{A}t} dt &= e^{\boldsymbol{A}h} \int_0^h e^{-\boldsymbol{A}(h-t)} dt \\ &= e^{\boldsymbol{A}h} \left(e^{-\boldsymbol{A}h} \cdot 0 + \int_0^h e^{-\boldsymbol{A}(h-t)} \boldsymbol{I} dt \right) \\ &= e^{\boldsymbol{A}h} \begin{bmatrix} \boldsymbol{I} & \boldsymbol{0} \end{bmatrix} \exp\left\{ \begin{bmatrix} -\boldsymbol{A} & \boldsymbol{I} \\ \boldsymbol{0} & \boldsymbol{0} \end{bmatrix} h \right\} \begin{bmatrix} \boldsymbol{0} \\ \boldsymbol{I} \end{bmatrix}. \end{aligned} \tag{8}$$

The objective function, $P(T)$, in the optimization problem in (5) is converted to discrete time by using the above, i.e.,

$$\begin{aligned} P(T) &= \sum_{k=0}^{N-1} \int_{kh}^{(k+1)h} \left(\boldsymbol{\Theta}(t)\boldsymbol{z}(t) + \tilde{\varphi}(t) \right) dt \\ &= \sum_{k=0}^{N-1} \boldsymbol{\Theta}_k \int_0^h \left(e^{\boldsymbol{A}t}\boldsymbol{z}_k + \int_0^t e^{\boldsymbol{A}(t-s)}\boldsymbol{B} ds \boldsymbol{u}_k \right) dt + h\tilde{\varphi}_k \\ &= \sum_{k=0}^{N-1} \boldsymbol{\Theta}_k \int_0^h \begin{bmatrix} \boldsymbol{I} & \boldsymbol{0} \end{bmatrix} e^{\tilde{\boldsymbol{A}}t} \begin{bmatrix} \boldsymbol{z}_k \\ \boldsymbol{u}_k \end{bmatrix} dt + h\tilde{\varphi}_k \\ &= \sum_{k=0}^{N-1} \boldsymbol{\Theta}_k \begin{bmatrix} \boldsymbol{I} & \boldsymbol{0} \end{bmatrix} e^{\tilde{\boldsymbol{A}}h} \begin{bmatrix} \boldsymbol{I} & \boldsymbol{0} \end{bmatrix} e^{\hat{\boldsymbol{A}}h} \begin{bmatrix} \boldsymbol{0} \\ \boldsymbol{I} \end{bmatrix} \begin{bmatrix} \boldsymbol{z}_k \\ \boldsymbol{u}_k \end{bmatrix} + h\tilde{\varphi}_k, \end{aligned}$$

where

$$\hat{\boldsymbol{A}} = \begin{bmatrix} -\tilde{\boldsymbol{A}} & \boldsymbol{I} \\ \boldsymbol{0} & \boldsymbol{0} \end{bmatrix}, \tilde{\boldsymbol{A}} = \begin{bmatrix} \boldsymbol{A} & \boldsymbol{B} \\ \boldsymbol{0} & \boldsymbol{0} \end{bmatrix}.$$

With the reference tracking disregarded and the growth of profit function as in (6), the optimization problem (5) can be reformulated as

$$\begin{array}{ll} \max\limits_{u_k \in U} & \sum\limits_{k=0}^{N-1} \boldsymbol{C}_k\boldsymbol{z}_k + \boldsymbol{D}_k\boldsymbol{u}_k + \boldsymbol{E}_k \\ \text{subject to} & \boldsymbol{z}_{k+1} = \boldsymbol{\Phi}\boldsymbol{z}_k + \boldsymbol{\Gamma}\boldsymbol{u}_k, \end{array} \tag{9}$$

where

$$\begin{aligned} C_k &= \boldsymbol{\Theta}_k \begin{bmatrix} \boldsymbol{I} & \boldsymbol{0} \end{bmatrix} e^{\tilde{\boldsymbol{A}}h} \begin{bmatrix} \boldsymbol{I} & \boldsymbol{0} \end{bmatrix} e^{\hat{\boldsymbol{A}}h} \begin{bmatrix} \boldsymbol{0} \\ \boldsymbol{I} \end{bmatrix} \begin{bmatrix} \boldsymbol{I} \\ \boldsymbol{0} \end{bmatrix}, \\ D_k &= \boldsymbol{\Theta}_k \begin{bmatrix} \boldsymbol{I} & \boldsymbol{0} \end{bmatrix} e^{\tilde{\boldsymbol{A}}h} \begin{bmatrix} \boldsymbol{I} & \boldsymbol{0} \end{bmatrix} e^{\hat{\boldsymbol{A}}h} \begin{bmatrix} \boldsymbol{0} \\ \boldsymbol{I} \end{bmatrix} \begin{bmatrix} \boldsymbol{0} \\ \boldsymbol{I} \end{bmatrix}, \\ E_k &= h\varphi_k,\ \boldsymbol{\Phi} = e^{\boldsymbol{A}(t_{k+1}-t_k)}, \text{ and } \boldsymbol{\Gamma} = \int_0^{t_{k+1}-t_k} e^{\boldsymbol{A}s} ds \boldsymbol{B}. \end{aligned}$$

[1] The prices used in this work corresponds to the market prices the 29th of June, 2008 and has been established using internal DONG Energy documents and the archive of power price at www.nordpool.dk, which is a marketplace for trading power contracts.

When considering the reference tracking different approaches can be used to formulate them. In this work two different methods are considered - briefly these are:

Quadratic: In this approach the tracking constraint is included in the profit function as a norm of the difference between the efficiency and the reference and thus penalizing deviations.

Side Constraint: In this approach the tracking is formulated as a constraint in the optimization such that the reference is followed within a reference band. This is implemented as additional side constraints to problem (9).

IV. Side Constraint

To include the reference tracking in problem (5) we introduce in this section a reference band with time dependent width, $\alpha(t)$, i.e., $\alpha(t)$ is the normed error at time t. In continuous time the reference band can be formulated as

$$h(\boldsymbol{z}(t), t) \geq 0, \tag{10}$$

where

$$h(\boldsymbol{z}(t), t) = \boldsymbol{\Upsilon}\boldsymbol{z}(t) + \boldsymbol{\psi}(t), \tag{11}$$

with

$$\boldsymbol{\Upsilon} = \begin{bmatrix} \boldsymbol{\gamma}^T\tilde{\boldsymbol{Q}} \\ -\boldsymbol{\gamma}^T\tilde{\boldsymbol{Q}} \end{bmatrix},$$
$$\boldsymbol{\psi}(t) = \begin{bmatrix} \boldsymbol{\gamma}^T\boldsymbol{b} - y_r(t) + \alpha \\ -\boldsymbol{\gamma}^T\boldsymbol{b} + y_r(t) + \alpha \end{bmatrix}.$$

By direct calculation the discrete time approximation then yields

$$\boldsymbol{\Psi}_l\boldsymbol{z}_k + \boldsymbol{\Pi}_l\boldsymbol{u}_k + \boldsymbol{\Omega}_{k,l} \geq 0$$

where for $l = 0, 1, 2, ..., L$

$$\boldsymbol{\Psi}_l = \boldsymbol{\Upsilon}e^{\boldsymbol{A}\frac{l-1}{L}h},$$
$$\boldsymbol{\Pi}_l = \boldsymbol{\Upsilon}\int_0^{\frac{l-1}{L}h} e^{\boldsymbol{A}(\frac{l-1}{L}h-s)}\boldsymbol{B}ds,$$
$$\boldsymbol{\Omega}_{k,l} = \boldsymbol{\psi}(\tfrac{l-1}{L}h + kh).$$

Note that the constraint is guaranteed to be satisfied L times between each sampling of the system in (9).

Hence the optimization problem (9) together with tracking constraint can be formulated as

$$\max_{\substack{u\,\in\,U \\ \alpha\,\geq\,0}} \sum_{k=0}^{N-1}\left(\boldsymbol{C}_k\boldsymbol{z}_k + \boldsymbol{D}_k\boldsymbol{u}_k + \boldsymbol{E}_k - W_k\alpha_k\right)$$
$$\text{subject to} \quad \boldsymbol{z}_{k+1} = \boldsymbol{\Phi}\boldsymbol{z}_k + \boldsymbol{\Gamma}\boldsymbol{u}_k,$$
$$\boldsymbol{\Psi}_l\boldsymbol{z}_k + \boldsymbol{\Pi}_l\boldsymbol{u}_k + \boldsymbol{\Omega}_{k,l} \geq 0.$$

Note that the tracking width α_k is included in the optimization problem, i.e., the tracking error is minimized as well.

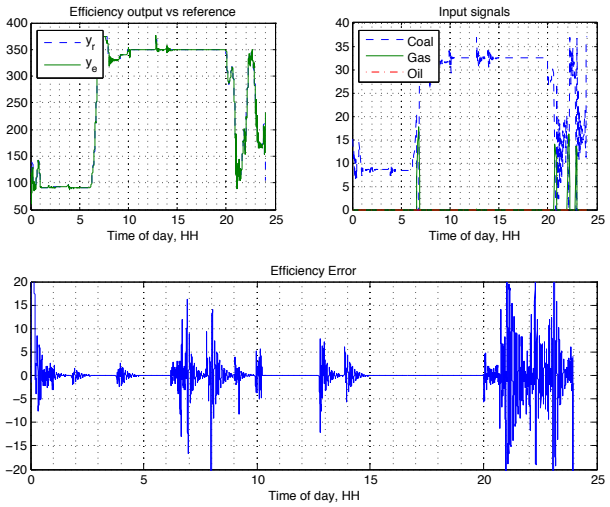


Fig. 1. Graphs of the efficiency output, input usage, and tracking error for the optimization problem with reference band tracking, $L = 1$.

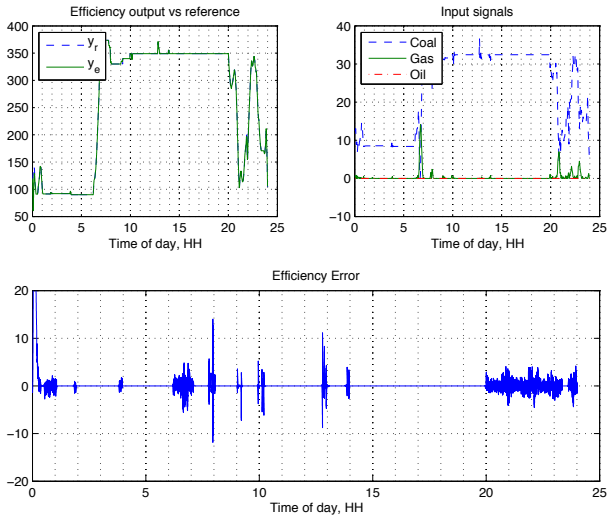


Fig. 2. Graphs of the efficiency output, input usage, and tracking error for the optimization problem with reference band tracking, $L = 5$.

The optimization problem above has been solved for $L = 1$ and $L = 5$, the results are depicted in Figure 2 and Figure 1. As seen in these figures, reference is tracked well with an significant error only present at times with large gradients in the reference signal. Furthermore, the tracking of the reference is considerably better when $L = 5$ as both the intensity and the value of the efficiency error is smaller.

V. Quadratic

In this section we include the reference tracking as a cost on the deviation from the reference. This is formulated as

$$Q(T) = \int_0^T -\beta_q \left\|\boldsymbol{\gamma}^T\boldsymbol{Q}\boldsymbol{z}(t) - y_r(t)\right\|^2 dt, \tag{12}$$

where $\|\cdot\|$ is the Euclidean norm. The tracking is included in the objective function as

$$\begin{aligned} P(T) &= \int_0^T g_p(\boldsymbol{z},t) - \beta_q \left\|\boldsymbol{\gamma}^T\boldsymbol{Q}\boldsymbol{z}(t) - y_r(t)\right\|^2 dt \\ &= \int_0^T \left(\underbrace{-\boldsymbol{z}(t)^T\boldsymbol{Q}\boldsymbol{z}(t)}_{P_2(T)} + \underbrace{2\boldsymbol{q}(t)^T\boldsymbol{z}(t) + \varphi(t)}_{P_1(T)} \right) dt, \end{aligned} \tag{13}$$

with

$$\begin{aligned} \boldsymbol{Q} &= \beta_q \tilde{\boldsymbol{Q}}^T \boldsymbol{\gamma}\boldsymbol{\gamma}^T \tilde{\boldsymbol{Q}} \\ \boldsymbol{q}(t)^T &= \frac{1}{2}\boldsymbol{\Theta}(t) + \beta_q y_r(t)\boldsymbol{\gamma}^T\tilde{\boldsymbol{Q}} \\ \varphi(t) &= \tilde{\varphi}(t) - \beta_q y_r(t)^2. \end{aligned}$$

As our maximization problem is formulated in discrete time we need to discretize (13). This is done in the sequel by apply (7) and (8).

$$\begin{aligned} P_1(T) &= \int_0^T \left(2\boldsymbol{q}(t)^T\boldsymbol{z}(t) + \varphi(t)\right) dt \\ &= \sum_{k=0}^{N-1} 2\boldsymbol{q}(t)^T \int_0^h \left(e^{\boldsymbol{A}t}\boldsymbol{z}_k + \int_0^t e^{\boldsymbol{A}(t-s)}\boldsymbol{B}ds\boldsymbol{u}_k \right) dt \\ &\quad + h\sum_{k=0}^{N-1}\varphi_k \\ &= \sum_{k=0}^{N-1} 2\boldsymbol{q}(t)^T \int_0^h \left[\begin{array}{cc} \boldsymbol{I} & \boldsymbol{0} \end{array}\right] e^{\tilde{\boldsymbol{A}}t} \left[\begin{array}{c} \boldsymbol{z}_k \\ \boldsymbol{u}_k \end{array}\right] dt + h\sum_{k=0}^{N-1}\varphi_k \\ &= \sum_{k=0}^{N-1} 2\boldsymbol{q}(t)^T e^{\tilde{\boldsymbol{A}}h} \left[\begin{array}{cc} \boldsymbol{I} & \boldsymbol{0} \end{array}\right] e^{\hat{\boldsymbol{A}}h} \left[\begin{array}{c} \boldsymbol{0} \\ \boldsymbol{I} \end{array}\right] \left[\begin{array}{c} \boldsymbol{z}_k \\ \boldsymbol{u}_k \end{array}\right] + h\sum_{k=0}^{N-1}\varphi_k \\ &= \sum_{k=0}^{N-1} \left(\boldsymbol{M}_z\boldsymbol{z}_k + \boldsymbol{M}_u\boldsymbol{u}_k + h\varphi_k\right), \end{aligned} \tag{14}$$

where

$$\begin{aligned} \boldsymbol{M}_z &= 2\boldsymbol{q}(t)^T e^{\tilde{\boldsymbol{A}}h} \left[\begin{array}{cc} \boldsymbol{I} & \boldsymbol{0} \end{array}\right] e^{\hat{\boldsymbol{A}}h} \left[\begin{array}{c} \boldsymbol{0} \\ \boldsymbol{I} \end{array}\right] \left[\begin{array}{c} \boldsymbol{I} \\ \boldsymbol{0} \end{array}\right] \\ \boldsymbol{M}_u &= 2\boldsymbol{q}(t)^T e^{\tilde{\boldsymbol{A}}h} \left[\begin{array}{cc} \boldsymbol{I} & \boldsymbol{0} \end{array}\right] e^{\hat{\boldsymbol{A}}h} \left[\begin{array}{c} \boldsymbol{0} \\ \boldsymbol{I} \end{array}\right] \left[\begin{array}{c} \boldsymbol{0} \\ \boldsymbol{I} \end{array}\right] \end{aligned}$$

with

$$\hat{\boldsymbol{A}} = \left[\begin{array}{cc} -\tilde{\boldsymbol{A}} & \boldsymbol{I} \\ \boldsymbol{0} & \boldsymbol{0} \end{array}\right], \quad \tilde{\boldsymbol{A}} = \left[\begin{array}{cc} \boldsymbol{A} & \boldsymbol{B} \\ \boldsymbol{0} & \boldsymbol{0} \end{array}\right],$$

and the matrices $\boldsymbol{I}$ and $\boldsymbol{0}$ of appropriate dimensions.

Now, the quadratic term is discretized by using (7)

$$\begin{aligned} P_2(T) &= -\boldsymbol{z}(t)^T\boldsymbol{Q}\boldsymbol{z}(t) \\ &= -\sum_{k=0}^{N-1}\int_0^h \left(\boldsymbol{z}_k^T e^{\boldsymbol{A}^T t} + \boldsymbol{u}_k^T\int_0^t \boldsymbol{B}^T e^{\boldsymbol{A}^T(t-s)}ds\right)\boldsymbol{Q} \\ &\qquad \left(e^{\boldsymbol{A}t}\boldsymbol{z}_k + \int_0^t e^{\boldsymbol{A}(t-s)}\boldsymbol{B}ds\boldsymbol{u}_k\right)dt \\ &= -\sum_{k=0}^{N-1}\int_0^h \left[\begin{array}{cc} \boldsymbol{z}_k^T & \boldsymbol{u}_k^T \end{array}\right] e^{\tilde{\boldsymbol{A}}^T t} \left[\begin{array}{c} \boldsymbol{I} \\ \boldsymbol{0} \end{array}\right] \boldsymbol{Q} \\ &\qquad \left[\begin{array}{cc} \boldsymbol{I} & \boldsymbol{0} \end{array}\right] e^{\tilde{\boldsymbol{A}}t} \left[\begin{array}{c} \boldsymbol{z}_k \\ \boldsymbol{u}_k \end{array}\right] dt \\ &= -\sum_{k=0}^{N-1} \left[\begin{array}{cc} \boldsymbol{z}_k^T & \boldsymbol{u}_k^T \end{array}\right] e^{\tilde{\boldsymbol{A}}^T h}\boldsymbol{Y}(h) e^{\tilde{\boldsymbol{A}}h} \left[\begin{array}{c} \boldsymbol{z}_k \\ \boldsymbol{u}_k \end{array}\right] \end{aligned} \tag{15}$$

where $\tilde{\boldsymbol{A}}$ is as above and

$$\begin{aligned} \boldsymbol{Y}(h) &= \int_0^h e^{-\tilde{\boldsymbol{A}}^T(h-t)}\bar{\boldsymbol{Q}}e^{-\tilde{\boldsymbol{A}}(h-t)}dt \\ \bar{\boldsymbol{Q}} &= \left[\begin{array}{c} \boldsymbol{I} \\ \boldsymbol{0} \end{array}\right] \boldsymbol{Q} \left[\begin{array}{cc} \boldsymbol{I} & \boldsymbol{0} \end{array}\right] \end{aligned} \tag{16}$$

The integral in (16) is on the form of the solution to a matrix differential equations which can be formulated as

$$\begin{aligned} \boldsymbol{Y}(h) &= \int_0^h e^{-\tilde{\boldsymbol{A}}^T(h-t)}\bar{\boldsymbol{Q}}e^{-\tilde{\boldsymbol{A}}(h-t)}dt \Rightarrow \\ -\frac{d}{dh}\boldsymbol{Y}(h) &= \tilde{\boldsymbol{A}}^T\boldsymbol{Y}(h) + \boldsymbol{Y}(h)\tilde{\boldsymbol{A}} - \bar{\boldsymbol{Q}}, \quad \boldsymbol{Y}(0) = \boldsymbol{0}. \end{aligned} \tag{17}$$

Using the $\mathrm{Vec}(\cdot)$ notation which is defined as

$$\mathrm{Vec}(\boldsymbol{P}) = \left[\begin{array}{c} p_1 \\ \vdots \\ p_n \end{array}\right], \tag{18}$$

where p_i is the columns of $\boldsymbol{P}$, it is possible to formulated (17) as

$$-\frac{d\mathrm{Vec}(\boldsymbol{Y}(h))}{dh} = \boldsymbol{F}\mathrm{Vec}(\boldsymbol{Y}(t)) - \mathrm{Vec}(\bar{\boldsymbol{Q}}) \tag{19}$$

where

$$\boldsymbol{F} = \left(\boldsymbol{I}\otimes\tilde{\boldsymbol{A}}^T + \tilde{\boldsymbol{A}}^T\otimes\boldsymbol{I}\right)$$

and $\otimes$ denotes the Kronecker product. By using the solution to standard vector differential equation and (8), the solution to (19) is given by

$$\begin{aligned} \mathrm{Vec}(\boldsymbol{Y}(h)) &= \int_0^h e^{\boldsymbol{F}(h-\tau)}d\tau\mathrm{Vec}(\bar{\boldsymbol{Q}}) \\ &= e^{\boldsymbol{F}h}\left[\begin{array}{cc} \boldsymbol{I} & \boldsymbol{0} \end{array}\right] e^{\hat{\boldsymbol{F}}h} \left[\begin{array}{c} \boldsymbol{0} \\ \boldsymbol{I} \end{array}\right] \mathrm{Vec}(\bar{\boldsymbol{Q}}) \\ &= e^{\boldsymbol{F}h}\tilde{\boldsymbol{F}}\mathrm{Vec}(\bar{\boldsymbol{Q}}), \end{aligned}$$

where

$$\tilde{\boldsymbol{F}} = \left[\begin{array}{cc} \boldsymbol{I} & \boldsymbol{0} \end{array}\right] e^{\hat{\boldsymbol{F}}h} \left[\begin{array}{c} \boldsymbol{0} \\ \boldsymbol{I} \end{array}\right], \quad \hat{\boldsymbol{F}} = \left[\begin{array}{cc} -\boldsymbol{F} & \boldsymbol{I} \\ \boldsymbol{0} & \boldsymbol{0} \end{array}\right].$$

That is (15) can be expressed as

$$P_2(T) = -\sum_{k=0}^{N-1} \left[\begin{array}{cc} \boldsymbol{z}_k^T & \boldsymbol{u}_k^T \end{array}\right] \left[\begin{array}{cc} \boldsymbol{N}_{zz} & \boldsymbol{N}_{zu} \\ \boldsymbol{N}_{uz} & \boldsymbol{N}_{uu} \end{array}\right] \left[\begin{array}{c} \boldsymbol{z}_k \\ \boldsymbol{u}_k \end{array}\right] \tag{20}$$

where

$$\boldsymbol{N}_{zz} = \left[\begin{array}{cc} \boldsymbol{I} & \boldsymbol{0} \end{array}\right] e^{\tilde{\boldsymbol{A}}^T h} \mathrm{Vec}^{-1}\left(e^{\boldsymbol{F}h}\tilde{\boldsymbol{F}}\mathrm{Vec}(\bar{\boldsymbol{Q}})\right) e^{\tilde{\boldsymbol{A}}h} \left[\begin{array}{c} \boldsymbol{I} \\ \boldsymbol{0} \end{array}\right]$$

$$\boldsymbol{N}_{zu} = \left[\begin{array}{cc} \boldsymbol{I} & \boldsymbol{0} \end{array}\right] e^{\tilde{\boldsymbol{A}}^T h} \mathrm{Vec}^{-1}\left(e^{\boldsymbol{F}h}\tilde{\boldsymbol{F}}\mathrm{Vec}(\bar{\boldsymbol{Q}})\right) e^{\tilde{\boldsymbol{A}}h} \left[\begin{array}{c} \boldsymbol{0} \\ \boldsymbol{I} \end{array}\right]$$

$$\boldsymbol{N}_{uz} = \left[\begin{array}{cc} \boldsymbol{0} & \boldsymbol{I} \end{array}\right] e^{\tilde{\boldsymbol{A}}^T h} \mathrm{Vec}^{-1}\left(e^{\boldsymbol{F}h}\tilde{\boldsymbol{F}}\mathrm{Vec}(\bar{\boldsymbol{Q}})\right) e^{\tilde{\boldsymbol{A}}h} \left[\begin{array}{c} \boldsymbol{I} \\ \boldsymbol{0} \end{array}\right]$$

$$\boldsymbol{N}_{uu} = \left[\begin{array}{cc} \boldsymbol{0} & \boldsymbol{I} \end{array}\right] e^{\tilde{\boldsymbol{A}}^T h} \mathrm{Vec}^{-1}\left(e^{\boldsymbol{F}h}\tilde{\boldsymbol{F}}\mathrm{Vec}(\bar{\boldsymbol{Q}})\right) e^{\tilde{\boldsymbol{A}}h} \left[\begin{array}{c} \boldsymbol{0} \\ \boldsymbol{I} \end{array}\right]$$

with matrices $\boldsymbol{I}$ and $\boldsymbol{0}$ of appropriate dimensions, and $\mathrm{Vec}^{-1}\left(e^{\boldsymbol{F}h}\tilde{\boldsymbol{F}}\mathrm{Vec}(\bar{\boldsymbol{Q}})\right)$, an $n \times n$ matrix, denoting the "inverse" of the Vec-operator in (18), i.e., reshaping the vector into a matrix.

Hence the optimization problem together with quadratic tracking error can be formulated as

$$\begin{aligned} &\max_{\substack{u \in U \\ \alpha \geq 0}} \sum_{k=0}^{N-1} C_k \\ \text{subject to} \quad &\boldsymbol{z}_{k+1} = \boldsymbol{\Phi}\boldsymbol{z}_k + \boldsymbol{\Gamma}\boldsymbol{u}_k, \end{aligned}$$

where

$$C_k = \left[\begin{array}{cc} \boldsymbol{z}_k^T & \boldsymbol{u}_k^T \end{array}\right] \boldsymbol{N} \left[\begin{array}{c} \boldsymbol{z}_k \\ \boldsymbol{u}_k \end{array}\right] + \boldsymbol{M}_z\boldsymbol{z}_k + \boldsymbol{M}_u\boldsymbol{u}_k + h\varphi_k,$$

with

$$\boldsymbol{N} = -\left[\begin{array}{cc} \boldsymbol{N}_{zz} & \boldsymbol{N}_{zu} \\ \boldsymbol{N}_{uz} & \boldsymbol{N}_{uu} \end{array}\right],$$

and the matrices $\boldsymbol{N}_{zz}$, $\boldsymbol{N}_{zu}$, $\boldsymbol{N}_{uz}$, $\boldsymbol{N}_{uu}$, $\boldsymbol{M}_z$, and $\boldsymbol{M}_u$ as given above.

The optimization problem above has been solved and the results are depicted in Figure 3. As seen in the figure

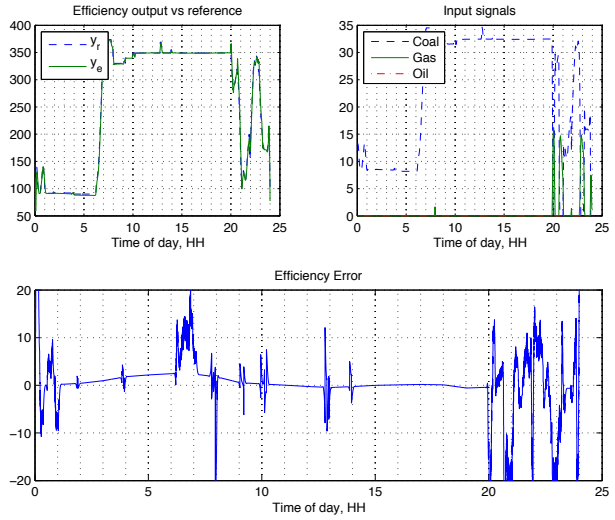


Fig. 3. Graphs of the efficiency output, input usage, and tracking error for the optimization problem with quadratic tracking error.

the reference is tracked with a small accumulated error caused by the quadratic error term in (12).

VI. Comparison of Optimization Methods

In this section we compare the two difference methods for solving the problem (9) with reference tracking.

Comparing the efficiency error of the three different methods it is noted that the mean of the error in the case of side constraints is less than the quadratic method. However, the fluctuations of the efficiency error when using side constraints are more frequent than the quadratic case.

The profit of the three different methods are almost identical and are therefore not included in this analysis. We note that this is also supported by the fact that the use of fuels in the three approaches are similar and by the fact that the efficiency error is small hence producing similar profits.

In table I the times for running the optimization are presented for the three solution strategies. Hence the quadratic method or side constraint method with $L = 1$ should be applied if only the optimization time is considered.

Method	Optimization Time	Solver
Side Constraint ($L = 5$)	896 s	SeDuMi
Side Constraint ($L = 1$)	168 s	SeDuMi
Quadratic	157 s	BPMPD

TABLE I

Comparison of optimization times between the three solution strategies.

As the profits of each of the methods are the same the choice of methods should be based on the need for computation time and requirements on tracking performance, which depend entire on the specific control problem.

To this end we remark that if continuous time is considered, the quadratic method has the advantage of only having the dynamical system as side constraint, which eases application of the Pontryagin maximum principle.

References

[1] D. Flynn, Ed., *Thermal Power Plant Simulation and Control*, ser. IEE Power & Energy Series. Michael Faraday House, Six Hills Way, Stenenage, Herts., SGI 2AY, United Kingdom: The Institute of Electrical Engineers, 2003, vol. 43.

[2] P. Andersen, J. D. Bendtsen, J. H. Mortensen, R. J. Nielsen, and T. S. Pedersen, "Observer-based fuel control using oxygen measurement - a study based on a first-principle model of a pulverized coal fired benson boiler." värmeforsk, Tech. Rep., 2005.

[3] K. Edlun, T. Mølbak, and J. D. Bendtsen, "Simple models for model-based portfolio load balancing controller synthesis," in *Proceeding of IFAC Symposium on Power Plants and Power Systems, Tampera, Finland*, 2009.

[4] A. Mas-Colell, M. D. Whinston, and J. R. Green, *Microeconomic Theory*. Oxford University Press, Inc., 1995.

[5] *Nord Pool*, http://www.nordpool.dk/, 2009, nord Pool is the Nordic electrical market, where power contracts are traded.

[6] S.-E. Fleten and T. K. Kristoffersen, "Short-term hydropower production planning by stochastic programming," *Computers & Operations Research*, vol. 35, no. 8, pp. 2656 – 2671, 2008, queues in Practice.

[7] M. Kragelund, R. Wisniewski, T. Mølbak, R. J. Nielsen, and K. Edlund, "On propagating requirements and selecting fuels for a benson boiler," in *Proceedings of the 17th IFAC World Congress, Seoul, South Korea*, 2008.

[8] M. Kragelund, J. Leth, and R. Wisniewski, "Selecting actuator configuration for a benson boiler: Production economics," in *Proceedings of the European Control Conference, Budapest, Hungary*, 2009.

[9] ——, "Optimal usage of coal, gas, and oil in a power plant," *IET Control Theory and Applications*, 2010, accepted for publication in Special issue on Advances in Complex Control Systems Theory and Applications.

[10] M. Kragelund, U. Jönsson, J. Leth, and R. Wisniewski, "Optimal production planning of a power plant," in *Proceedings of the International Conference on Control and Automation, Christchurch*, 2009.

[11] K. J. Åström and B. Wittenmark, *Computer-Controlled Systems - Theory and Design*. Prentice-Hall International, 1990.

2 Modeling and Control of an Experimental pH Neutralization Plant using Neural Networks based Approximate Predictive Control

Modeling and Control of an Experimental pH Neutralization Plant using Neural Networks based Approximate Predictive Control

Ireneus Wior
Institute of Automation Technology
Helmut-Schmidt-University, Germany
Email: ireneuswior@arcor.de

Sudchai Boonto
Hossam Seddik Abbas
and Herbert Werner
Institute of Control Systems
Hamburg University of Technology, Germany

***Abstract*—A nonlinear experimental pH neutralization plant is controlled using a neural networks based Approximate Predictive Control (APC) strategy. First a closed-loop identification is performed, further, using neural networks, a black-box modeling of the experimental plant is conducted. Then the approximate predictive controller is realized, where a linear model of the plant is extracted at each sampling period from the neural network model. This strategy is used to control the experimental neutralization plant for set point tracking and disturbance rejection.**

***Index Terms*—pH Neutralization Plant, Neural Networks, Approximate Predictive Control**

I. Introduction

Monitoring and controlling the pH level is often performed in many chemical, industrial processes. It is important to improve the productivity and at the same time the robustness of these processes.

A PID controller is often used to deal with this process; however, it can only react to changes in a reference signal. On the other hand, a Model Predictive Control (MPC) approach is proactive and makes use of the information of the future reference signal which is usually known beforehand in a pH neutralization process. Although the MPC approach can not follow a step function directly it can follow it much better than a PID (justifiably assuming that the maximal possible slope of the change is the same). Furthermore a MPC approach is more sophisticated than a PID in terms of handling input and output constraints, as well as dealing with difficult system behaviors like high nonlinearity and long time delays, see e.g. [1].

Nonlinear Model Predictive Control (NMPC) is a well-established research approach to deal with nonlinear plants. Currently the NMPC is limited to processes with relatively slow dynamics due to the usage of nonlinear optimization approaches. Different techniques have been proposed to deal with this problem, see e.g. [2] and [3]. One of these techniques is the Approximate Predictive Control (APC) which uses a linearized model of the plant at each sampling period [4]. By this way, only a linear optimization problem has to be performed every sampling instant, this reduces the computational load and enables to deal with faster processes.

The APC approach is already known since more than ten years and some simulation studies were introduced in [5] to control gas turbine engines and in [4] for a pneumatic servomechanism. Although there is a vast literature on MPC in connection with neural networks there are not many applications of APC approaches based on neural network models used on real problem instances. One of the exceptions is the experimental 3-DOF Helicopter presented in [6].

To deal with the pH neutralization process in [7] a PID controller based on a neural network model is presented, which uses a genetic algorithm to tune the parameters of the PID controller offline on the nonlinear neural network model. In [8] on this problem an adaptive nonlinear control strategy is used.

Fig. 1. The experimental pH neutralization plant.

In this paper an APC strategy is used to control the experimental pH neutralization plant shown in Figure 1. We consider the set point tracking and disturbance rejection problem. All

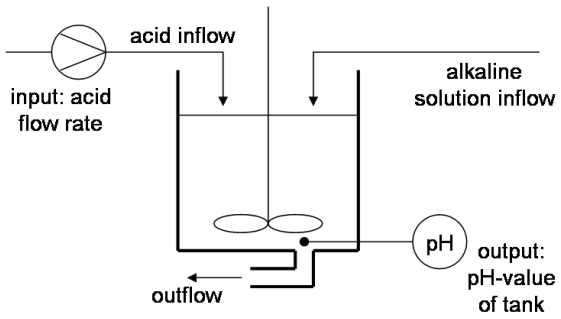


Fig. 2. Schema of the problem in its SISO arrangement.

these in order to assist the capability of APC to control such highly nonlinear process.

The rest of the paper is structured as follows: Section II describes the system under consideration. In Section III a system identification of the experimental pH neutralization plant is performed. Here a data set is collected in closed-loop and based on which a neural network is trained to represent the experimental plant. A short introduction to the APC strategy is given in Section IV. Section V presents the experimental results in terms of a set point tracking and a disturbance rejection problems. Finally in Section VI some conclusions are drawn.

II. Problem description

The pH neutralization process considered in this work is technically realized in a mixing tank with two input streams and one output stream. Figure 1 and Figure 2 show the experimental plant and a draft of it, respectively. The cylindrical tank is initially filled to three-fourths of its volume with water and the mixer is arranged in the lower fourth of the tank.

Separate control loops, one for the temperature and one for the liquid level, are used for holding the environmental conditions approximately constant. These controllers are simple On/Off-Controllers and as such are independent from the control method for the pH neutralization process.

One of the input streams is an alkaline solution (NaOH) and has a constant flow rate as well as a constant pH-value. The second input stream is acid (HCl) with a constant pH-value but its flow rate is manipulated to control the pH-value in the tank. The output stream is controlled in the mentioned separated On/Off-Control loop and hence the outflow is discrete depending on the liquid level within the tank. Finally a pH sensor is attached near the bottom of the tank precisely above the opening for the output stream. The reference signal which is a desired pH-value in the tank is known beforehand.

The whole system can be formulated as a SISO system: the pH-value of the liquid in the tank is the output and the acid flow rate is the input to this system. The right acid flow rate results in the desired pH-value within the tank. Insufficient acid results in an excessively alkaline pH-value; conversely, excessive acid inflow leads to an exceedingly acidic pH value.

In the following the nonlinearities of the experimental pH neutralization process are summarized:

- The neutralization process proceeds nonlinearly and has a high sensitivity around the pH-value of 7.
- The neutralization process has two regions of saturation: one in the very acidic (pH-value < 5 approximately) and one in the very alkaline region (pH-value > 9 approximately). If one of the regions is reached, it is very difficult to lead the pH-value out of the saturation.
- Although a mixer blends the liquids in the tank, a continuous homogenous distribution can not be reached immediately. One reason for this is the positions of the mixer relative to the location of the inflow stream, with the first being placed at the bottom of the tank, while the latter is found at the top.
- As a consequence of inhomogeneous liquid distribution, the liquid level, as well as other effects, the whole system has a noticeable time delay.

Furthermore the experimental build-up has some limitations:

- The available storage volume for the acid solution is limited and leads to a limitation of the measurement duration. Hence the data set which can be collected is relatively small.
- The pH sensors used have a measuring range from 0 to 14 pH, a smallest measuring range of 0.5 pH and an accuracy of $\pm$ 0.2%. This accuracy of the pH sensors defines the highest accuracy for the control.

III. System identification

To handle the nonlinear character of the plant a *black-box modeling* method is used. First the data has to be generated with which a neural network is then trained.

A. Data generation

For black-box modeling a set of input and output data must contain all important information about the behavior of the plant. To get all important information of a nonlinear system, the whole range of amplitudes and frequencies must be stimulated within which the plant shall be operated. The resulting data is a set of data input u_k and output y_k of the experimental plant with N being the number of samples k: $Z^N = \{u_k, y_k \mid k = 1, 2, \ldots, N\}$. Because of the saturation regions of the neutralization process the amplitude range of the pH-value which has to be covered is from around 5 to around 9.

With the *relay feedback method* [9] the critical frequency f_c; with different step responses the rise time t_r and finally using (1) the sampling frequency f_s and hence the sampling time T_s are determined as $f_c = \frac{1}{60}$Hz, $t_r = 90$sec and $T_s = 9$sec, respectively.

$$f_s = (5 \sim 10) \cdot f_c \quad and \quad f_s = (5 \sim 10) \cdot \frac{1}{t_r} \qquad (1)$$

A *multisine* signal [10] is used to excite the system. This is a periodic non-binary multifrequency signal given as:

$$u(t) = \sum_{i=1}^{n_s} A_i \cdot cos(\omega_i \cdot t), \tag{2}$$

where A_i and ω_i are the i-th amplitude and frequency of the multisine. With a multisine a desired frequency spectrum with constant amplitudes in a desired frequency range can be designed easily. Additionally a relatively small crest factor can be achieved (hence it has a good signal-to-noise ratio). Following [11] the minimum number of samples N and the minimum number of different frequencies n_s of the multisine can be computed by:

$$N \geq \frac{2 \cdot \pi \cdot \beta_s \cdot \tau_{dom}}{T}, \tag{3}$$

$$n_s \geq \frac{N \cdot T_s \cdot \alpha_s}{2 \cdot \pi \cdot \tau_{dom}}, \tag{4}$$

where τ_{dom} is the dominant plant time constant, β_s specifies how much low-frequency information will be in the signal and here it is chosen as $\beta_s = 3$ to get low-frequency information. The constant α_s denotes how much faster the closed-loop response is expected to be in comparison with the open-loop one, it is chosen as $\alpha_s = 1$. In addition, N and n_s are chosen as 350 and 20, respectively, which fulfil (3) and (4).

Figure 3 shows a typical spectrum of an input signal for the identification purpose, where f_n is the Nyquist frequency, f_b is the bandwidth of the closed loop system which is taken as $f_b = \frac{1}{60}$Hz with $f_b = \alpha_s \cdot f_c$. The low frequency part up to f_b stimulates the range in which the plant shall be operated. The amplitude of the high frequency part from f_b to f_n is only half of the amplitude of the low frequency part. Therefor, the high frequency noise is not significantly amplified.

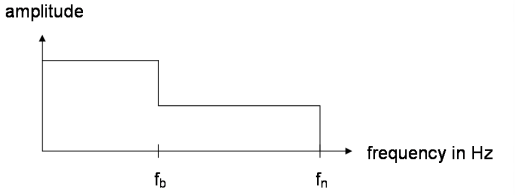


Fig. 3. The frequency spectrum of the input signal.

The closed-loop shown in Figure 4 is used to generate the data set for identification. By using the feedback control scheme one can force the output signal to get out of the saturation. This counteracts the problem of getting stuck in the saturation which otherwise occurs with the open-loop approach. A proportional controller is used in the closed-loop and the input signal is added just after the proportional controller. It is known that closed-loop identification based on a direct approach [10] is sensitive to noise since the noise of the input to the plant is correlated with the noise of the output; however, the high signal-to-noise ratio allows to assume that the amount of the noise in the output signal is neglectable.

The data set generated in this way is shown in Figure 5 and will be used in the following section for the training of the neural network.

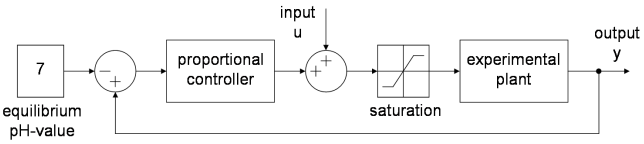


Fig. 4. Closed loop structure for the identification of the data set.

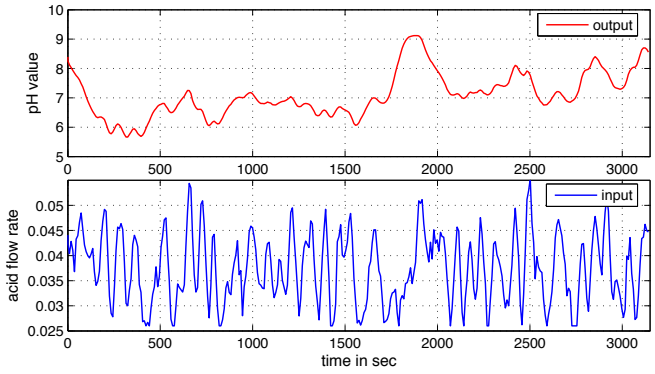


Fig. 5. Input and output signals of the training data set.

B. System identification with neural networks

A neural network is trained to capture the nonlinear behavior of the plant. The structure of the neural network is chosen to correspond to an ARX (AutoRegressive with eXogenous input) model structure in linear systems [10]. We refer to a neural network with this structure as a NNARX model structure, see Figure 6.

The input vector φ of the neural network consists of the past n output signals y_{k-1} until y_{k-n} and of the past m input signals u_{k-d} until $u_{k-d-m+1}$ which are shifted by the delay d. The output of the neural network is $\hat{y}_k$ which is a prediction of the plant's output at instant k. A *multilayer preceptron* neural network type is used [4], with two layers, p neurons and which is described as:

$$\hat{y} = f^2(W^2 f^1(W^1 \varphi + \omega_0^1) + \omega_0^2), \tag{5}$$

where f^1 and f^2 are the tangent hyperbolic and linear functions, respectively, W^1 and W^2 are matrices containing the network weights and ω_0^1 as well as ω_0^2 are the weights of the biases. θ in Figure 6 contains all weights, i.e. it includes the weights W^1, W^2, ω_0^1 and ω_0^2.

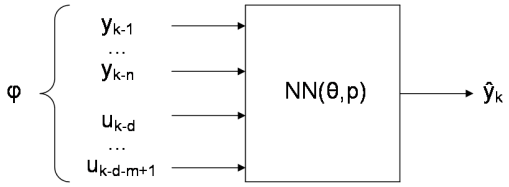


Fig. 6. NNARX model structure.

To train the network, i.e. to find the weights, the *Levenberg-Marquardt Backpropagation algorithm* [4] is used. The method

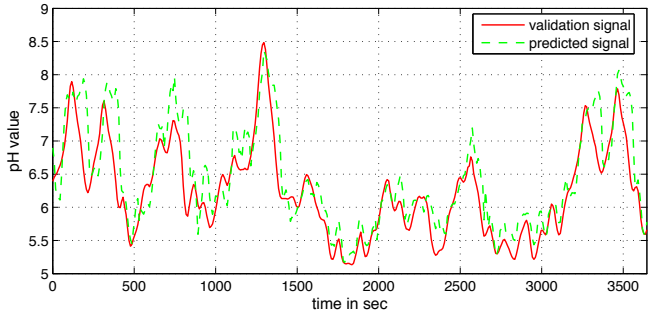


Fig. 7. Validation signal (solid) and 10-step-ahead prediction (dashed) of a NNARX model with $p = 11$, $n = 11$, $m = 10$ and $d = 1$.

seeks to minimize the sum of the mean squared prediction errors given as:

$$V_N(\theta, Z^N, \alpha) = \frac{1}{2N}\sum_{k=1}^{N}((y_k - \hat{y}_k(\theta))^2 + \frac{1}{2N}\theta^T \cdot \alpha I \cdot \theta). \quad (6)$$

During the training of the neural network an undesired effect may occur, which is known as *overfitting* [4], [12]. In overfitting, the neural network is not only trained on the plant dynamics but also on the plant disturbance. In order to deal with this, two methods can be used: training with a *regularization term* and *pruning* [4], [12]. Both methods are used in this paper. The regularization term α can be found in (6) and it is tuned to be as $\alpha = 10^{-3}$.

To implement these methods the *Neural Network Based System Identification TOOLBOX* [13] is used. It contains algorithms for the training and the validation of multilayer perceptron neural networks together with methods for pruning and the regularization term.

Using the data generated in the closed-loop, a NNARX model of the experimental neutralization plant is found. The result is a neural network which has $p = 11$ neurons, uses $n = 11$ past outputs as well as $m = 10$ past inputs and has a delay of $d = 1$. The number of past inputs m and past outputs n was determined with an order index criterion based on Liptschitz quotients [4]. The 10-step ahead prediction with the NNARX model in comparison to the validation signal can be seen in Figure 7. It has to be noted that the data set used to train the neural network and the data set used for validation are two different ones. Since the storage volume of the acid solution is limited the measurement period is also limited. This may reduce the quality of the nonlinear model; however, as shown in Figure 7 the plant behavior has been identified with satisfactory in the 10-step ahead prediction. In the following sections this neural network model is used for the APC controller as well as for tuning the controller in a simulation build-up.

IV. Approximate Predictive Control

The main concept behind common predictive control strategies is to predict the future outcome of different plant inputs and to choose the best out of these. Its calculations are relatively time consuming, this being its main disadvantage.

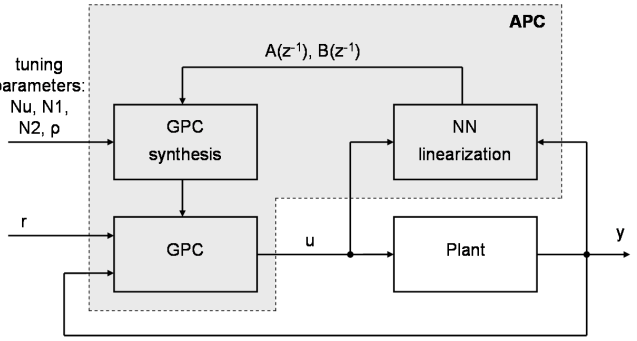


Fig. 8. Block structure of Approximate predictive control.

The minimization problem:

$$\min_{\tilde{U}_k} J_k = \min_{\tilde{U}_k}(\sum_{i=N_1}^{N_2}(r_{k+i} - \hat{y}_{k+i})^2 + \rho\sum_{j=1}^{N_u}\Delta u_{k+j-1}^2), \quad (7)$$

where J_k is the cost function, r_{k+i} is the known future reference signal and ρ is a factor which penalizes the influence of the input signals on the cost function has to be solved at each instant k. Moreover $\tilde{U}_k$ is a vector with the most recent control input changes given as:

$$\tilde{U}_k = [\Delta u_k \quad \Delta u_{k+1} \quad \dots \quad \Delta u_{k+Nu-1}]^T, \quad (8)$$

where $\Delta = 1 - z^{-1}$ with z^{-1} is a time delay operator (i.e. $z^{-p}u_k = u_{k-p}$). At each instant only the first computed input change Δu_k is applied to the plant and then the whole computation is repeated for the next instant.

To solve this minimization problem the predicted outputs $\hat{y}_{k+i}$ within the fixed prediction horizon have to be determined. To reduce the calculation time requirements the *General Predictive Control* (GPC) approaches use a linear model to predict the future outputs [14], [15]. This results in a linear optimization problem with a new linear model for each sampling period. *Approximate Predictive Control* (APC) is a special case of the GPC approach, where the linear model is extracted from a neural network. This is known as *instantaneous linearization* [4]. In Figure 8 the block structure of the APC is shown.

A detailed introduction of the APC can be found in [4], and a toolkit which uses these formulas is implemented in [16].

V. Experimental Study

In this section the experimental results of the APC on the pH neutralization plant are presented. A set point tracking problem as well as a disturbance rejection problem are considered. It is difficult to adjust the parameter values of the controller directly in the real experimental plant because of its nonlinear behavior, the time intensive preparations to run the plant and the long measurement duration to obtain sufficient measurements. Therefore, first the parameter values of the APC are adjusted off-line with the neural network model identified earlier, then the APC is tried on the experimental plant.

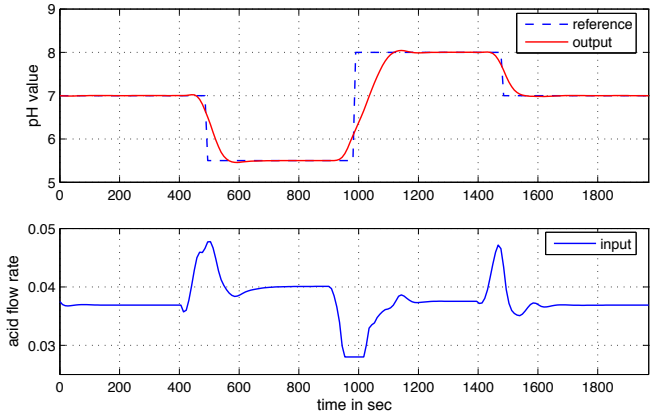


Fig. 9. Simulation results of a set point tracking problem with APC ($N1 = 1$, $N2 = 10$, $Nu = 2$ and $\rho = 2000$).

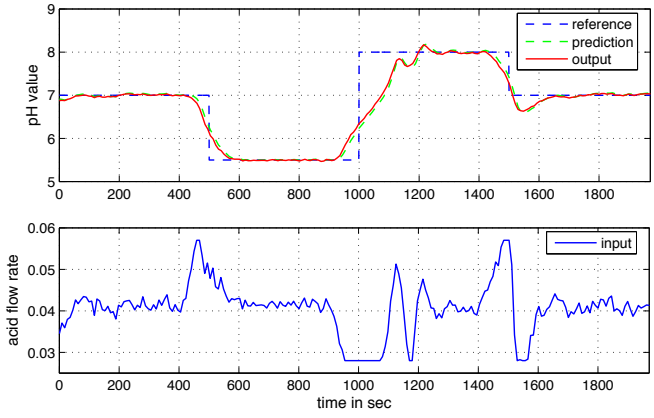


Fig. 10. Experimental results of a set point tracking problem with APC ($N1 = 1$, $N2 = 10$, $Nu = 2$ and $\rho = 20000$).

A. *Off-line parameter values adjustment*

The neural network model found in Section III-B which provides the results shown in Figure 7 is used to tune the APC off-line. The model is used twofold: first it is applied inside the APC structure (for this purpose the neural network model is constructed as shown in Figure 8), and second it simulates the plant that has to be controlled.

The parameter values of the APC are adjusted as follows: $N_1 = 1$ is fixed and equals the delay of the system. N_2 is selected as $\frac{t_r}{T_s} = 10$ so that the prediction horizon covers at least the rise time of the plant; $N_u = 2$ is chosen relatively small in comparison with N_2. It has been observed that the choice of N_2 and N_u is mostly unproblematic and gives good results for different values. The value of ρ in (7), which penalizes the control signal, should be carefully tuned. With $\rho = 2000$ a satisfactory tracking capability has been achieved in simulation.

Figure 9 shows the simulation results of the set point tracking problem. The reference signal is a three level signal that changes each 500 seconds. The APC produces reasonable control inputs and the output tracks the reference signal in a satisfactory manner. Finally, in Figure 9 the proactive characteristic of the APC can be observed because the control action begins earlier than the change in the reference. Next, the above parameter values are used with the real experimental plant.

B. *Experimental results*

The same obtained controller parameters have been utilized when the APC is tested on the real experimental pH neutralization plant. However, to improve the tracking capability and to reduce some oscillations which appear during the implementation, the value of ρ has been further tuned online, and it turns out that its best value is $\rho = 20000$.

The resulted measurement on the set point tracking problem is shown in Figure 10. The first change from pH-value 7 to 5.5 is unproblematic and relatively well done. The changes from pH 5.5 to 8 and then back to 7 are not ideal, but the control signal touches its limits, in particular, the lower one. Overall the results are reasonable and almost the same as the simulated ones.

From Figure 10 it can be also seen that the prediction is close to the real output, which shows that the model can capture the dynamics of the plant very well.

Finally the performance of the APC on a disturbance rejection problem is considered. The same parameter values as in the set point tracking problem are used. The task for the controller is to hold a constant pH-value equal 6, while some unmeasured disturbances are acting on the alkaline input stream. In Figure 11 the results are presented. The first disturbance is done by increasing the base valve opening for one sampling period from 0.2 to 0.4, which is equivalent to a three times higher alkaline flow rate. The second disturbance is obtained by reducing the valve opening from 0.2 to 0 for two sampling periods. It can be observed that the controller directly reacts with a change of the acid inflow when a deviation in the pH value occurs. Furthermore, it can be seen that the disturbance can not bring the pH-value far from the reference.

VI. Conclusion

In this work an approximate predictive control strategy for an experimental pH neutralization plant has been carried out. In closed-loop, with a multisine input signal, an identification data set has been gathered. A multilayer preceptron network with a NNARX model structure has been trained. Based on the trained neural network model, an APC has been off-line tuned and then implemented on the experimental plant.

The experimental results of a set point tracking and a disturbance rejection problems have demonstrated the capability of the APC to control the experimental pH neutralization plant successfully.

References

[1] J.M. Maciejowski: *Predictive control with constraints*. Prentice-Hall, Pearson Education Limited, Harlow, UK, 2002.

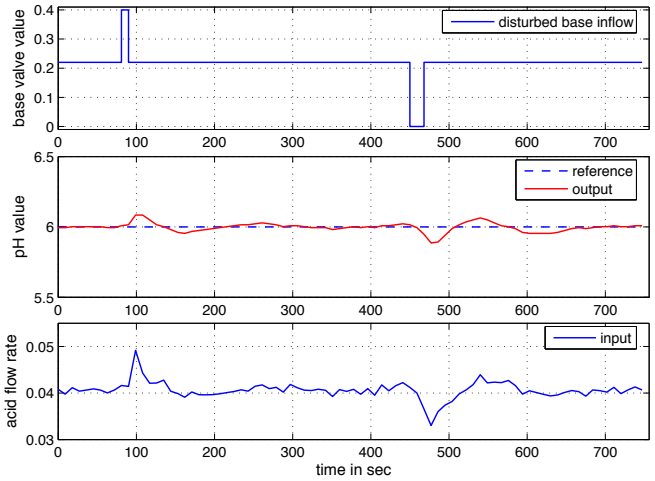


Fig. 11. Experimental results of a disturbance rejection problem with APC ($N1 = 1$, $N2 = 10$, $Nu = 2$ and $\rho = 20000$).

[2] V.M. Zavala, C.D. Laird, L.T. Biegler: *Fast implementations and rigorous models: Can both be accommodated in NMPC?*. International Journal of Robust and Nonlinear Control, Vol. 18, No. 8, pp. 800-815, 2008.

[3] J.H. Lee: *Modeling and Identification for Nonlinear Model Predictive Control: Requirements, Current Status and Future Research Needs*. Nonlinear Model Predictive Control, F. Allgöwer and A. Zheng (Eds.), Progress in Systems and Control Theory, Vol. 26, Birkhauser Verlag, Basel 2000.

[4] M. Norgaard, O. Ravn, N. Poulsen, and L. Hansen: *Neural Networks for Modelling and Control of Dynamic Systems*. Springer-Verlag, London, UK, 2000.

[5] J. Mu, D. Rees: *Approximate model predictive control for gas turbine engines*. In Proc. American Control Conference, 2004, pp. 5704-5709.

[6] J. Witt, S. Boonto, H. Werner: *Approximate model predictive control of a 3-DOF helicopter*. In Proc. IEEE Conf. Decision and Control, 2007, pp. 4501-4506.

[7] A. Popov, A. Farag and H. Werner: *Tuning of a PID controller using a multiobjective optimization technique applied to a neutralization plant*. In Proc. IEEE Conf. Decision Control, 2005, pp. 71397143.

[8] M.A. Henson, D.E. Seborg: *Adaptive nonlinear control of a pH neutralization process*. In IEEE Transactions on Control Systems Technology, Vol. 2, No. 3, pp. 169-182, 1994.

[9] K. J. Aström and B. Wittenmark: *Adaptive Control*. Addison-Wesley Series in Electrical Engineering: Control Engineering. Addison-Wesley Publishing Co., 2nd edition edition, 1995.

[10] L. Ljung: *System Identification, Theory for the User*. Prentice Hall Information and System Sciences Series. Prentice-Hall Inc. USA, 2nd edition edition, 1999.

[11] D. E. Rivera, S. V. Gaikwad, and X. Chen: *A demonstration prototype for control-relevant identification*. Technical report, Dept. of Chemical, Bio and Materials Engineering, Arizona State University, 1994.

[12] I.R. Wior: *Modeling and Control for Titration and Neutralizatio (TINA) Plant*. Institute of Control Systems, Hamburg University of Technology, Bachelor Thesis, 2006.

[13] M. Norgaard: *Neural network based control system design toolbox*, ver. 2. Technical Report 00-E-891, Department of Automation, Technical University of Denmark, 2000.

[14] D.W. Clarke, C. Mothadi, P.S. Tuffs: *Generalized predictive control - Part I. The basic algorithm*. Automatica, Vol. 23, No. 2, pp. 137-148, 1987.

[15] D.W. Clarke, C. Mothadi, P.S. Tuffs: *Generalized predictive control - Part II. Extensions and interpretations*. Automatica, Vol. 23, No. 2, pp. 149-160, 1987.

[16] M. Norgaard: *Neural network based control system design toolkit*, ver. 2. Technical Report 00-E-892, Department of Automation, Technical University of Denmark, 2000.

3 Sensitivity Analysis of the LMI-based H-inf Control Problem

Sensitivity Analysis of the LMI-based $\mathcal{H}_\infty$ Control Problem

A. S. Yonchev*, P. Hr. Petkov*, N. D. Christov† and M. M. Konstantinov‡
*Department of Systems and Control, Technical University of Sofia
1000 Sofia, Bulgaria; Email: ajonchev@mail.bg, php@tu-sofia.bg
†Laboratory of Automatics, Computer Engineering and Signal Processing
Lille University of Science and Technology, 59655 Villeneuve d'Ascq
France; Email: Nicolai.Christov@univ-lille1.fr
‡Department of Mathematics, University of Architecture, Civil Engineering
and Geodesy, 1046 Sofia, Bulgaria; Email: mmk_fte@uacg.bg

***Abstract*—Local perturbation bounds are obtained for the continuous-time $\mathcal{H}_\infty$ control problem based on linear matrix inequalities (LMI). The sensitivity analysis of the perturbed LMI is done by introducing a suitable slightly perturbed right-hand part. This approach leads to tight, condition number based perturbation bounds for the LMI solutions to the $\mathcal{H}_\infty$ control problem.**

I. Introduction

In the last decade a number of papers have been published on the sensitivity of the $\mathcal{H}_\infty$ control problem [2]. These papers, however, consider exclusively the case of the Riccati-based $\mathcal{H}_\infty$ control problem. In contrast, in this paper we study the sensitivity of the LMI-based $\mathcal{H}_\infty$ control problem. We propose a new approach to the perturbation analysis of this problem via introducing a suitable right hand part in the considered matrix inequalities. Using this new perturbation technique we obtain local perturbation bounds for the the continuous-time LMI-based $\mathcal{H}_\infty$ control problem in terms of condition numbers with respect to the perturbations in the data.

We use the following notations: $\mathbb{R}^{m\times n}$ – the space of real $m \times n$ matrices; $\mathbb{R}^n = \mathbb{R}^{n\times 1}$; I_n – the identity $n \times n$ matrix; e_n – the unit $n \times 1$ vector; $M^\top$ – the transpose of M; $M^\dagger$ – the pseudo inverse of M; $\|M\|_2 = \sigma_{\max}(M)$ – the spectral norm of M, where $\sigma_{\max}(M)$ is the maximum singular value of M; $\|M\|_{\rm F} = \sqrt{\rm tr(M^\top M)}$ – the Frobenius norm of M; $\|M\|_\infty := \sup_{{\rm Re}\ s\geq 0}\|M(s)\|_2$; $\|.\|$ is any of the above norms; $\rm vec(M) \in \mathbb{R}^{mn}$ – the column-wise vector representation of $M \in \mathbb{R}^{m\times n}$; $\Pi_{m,n} \in \mathbb{R}^{mn\times mn}$ – the vec-permutation matrix, such that $\rm vec(M^\top) = \Pi_{m,n} vec(M)$; $M \otimes P$ – the Kroneker product of the matrices M and P; $\rm vec(MXP) = (P^\top \otimes M) vec(X)$ – column-wise vector representation of the multiplication MXP. The notation ":=" stands for "equal by definition".

The paper is organized as follows. In Section II we shortly present the problem setup and objective. Section III describes the performed linear sensitivity analysis of the LMI based $\mathcal{H}_\infty$ control problem. Section IV presents a numerical example before we conclude in Section V with some final remarks.

II. Problem Statement

Consider the linear continuous-time system

$$\begin{aligned} \dot{x}(t) &= Ax(t) + B_1 w(t) + B_2 u(t) \\ z(t) &= C_1 x(t) + D_{11} w(t) + D_{12} u(t) \\ y(t) &= C_2 x(t) + D_{21} w(t) \end{aligned} \tag{1}$$

where $x(t) \in \mathbb{R}^n$, $u(t) \in \mathbb{R}^m$, $y(t) \in \mathbb{R}^r$ and $z(t) \in \mathbb{R}^p$ are the system state, input, output and performance vectors respectively, $w(t) \in \mathbb{R}^l$ is the disturbance and A, B_1, B_2, C_1, C_2, D_{11}, D_{12}, D_{21} are constant matrices of compatible size.

The suboptimal $\mathcal{H}_\infty$ control problem consists in finding a control law $u(t)$ which leads to a bounded $\mathcal{H}_\infty$-norm of the transfer function matrix $T_{zw}(s)$ from $w(t)$ to $z(t)$:

$$\|T_{zw}(s)\|_\infty < \gamma, \quad \gamma > 0. \tag{2}$$

In the optimal $\mathcal{H}_\infty$ control problem one tries to find the infimum of γ (further denoted by γ_{opt}) which satisfies (2). The solution of the optimal $\mathcal{H}_\infty$ control problem corresponds to the best disturbance attenuation at the performance vector of the closed-loop system.

The $\mathcal{H}_\infty$ control problem (1), (2), is solvable if and only if there exist two symmetric matrices $R, S \in \mathbb{R}^{n\times n}$ satisfying the following system of LMI [4] :

$$\begin{bmatrix} \mathcal{N}_{12} & \vdots & 0 \\ \cdots & \cdots & \cdots \\ 0 & \vdots & I \end{bmatrix}^\top \begin{bmatrix} AR + RA^\top & RC_1{}^\top & \vdots & B_1 \\ C_1 R & -\gamma I & \vdots & D_{11} \\ \cdots & \cdots & \cdots & \cdots \\ B_1{}^\top & D_{11}{}^\top & \vdots & -\gamma I \end{bmatrix}$$

$$\times \begin{bmatrix} \mathcal{N}_{12} & \vdots & 0 \\ \cdots & \cdots & \cdots \\ 0 & \vdots & I \end{bmatrix} < 0 \tag{3}$$

$$\left[\begin{array}{ccc}\mathcal{N}_{21} & \vdots & 0\\ \cdots & \cdots & \cdots\\ 0 & \vdots & I\end{array}\right]^{\top}\left[\begin{array}{cccc}A^{\top}S+SA & SB_1 & \vdots & C_1{}^{\top}\\ B_1{}^{\top}S & -\gamma I & \vdots & D_{11}{}^{\top}\\ \cdots & \cdots & \cdots & \cdots\\ C_1 & D_{11} & \vdots & -\gamma I\end{array}\right]$$

$$\times\left[\begin{array}{ccc}\mathcal{N}_{21} & \vdots & 0\\ \cdots & \cdots & \cdots\\ 0 & \vdots & I\end{array}\right]<0, \tag{4}$$

$$\left[\begin{array}{cc}R & I\\ I & S\end{array}\right]>0 \tag{5}$$

where $\mathcal{N}_{12}$ and $\mathcal{N}_{21}$ are the orthonormal bases of the null spaces of $\left[\begin{array}{cc} B_2{}^{\top} & D_{12}{}^{\top}\end{array}\right]$ and $\left[\begin{array}{cc} C_2 & D_{21}\end{array}\right]$, respectively.

Computing solutions (R,S) of the LMI system (3)-(5) is a convex optimization problem. The sensitivity of the LMI under consideration, subject to variations in the system data, may affect the accuracy of the matrices R and S and hence the accuracy of controller matrices. It is not clear up to the moment how LMI sensitivity is connected to the sensitivity of the given $\mathcal{H}_\infty$ suboptimal problem.

In what follows, we assume that $\gamma_{\rm opt}$ is determined and present a sensitivity analysis of the optimal $\mathcal{H}_\infty$ control problem based on the LMI (3)-(5).

Suppose that the matrices $A,\ \ldots,\ D_{21}$ and the quantity γ in (3), (4) are subject to perturbations $\Delta A,\ \ldots,\ \Delta\gamma_{\rm opt}$ and denote by $R^*+\Delta R$, $S^*+\Delta S$ the solution of the perturbed LMI system. The sensitivity analysis of the $\mathcal{H}_\infty$ control problem is aimed at determining bounds for ΔR and ΔS near the optimal value of γ, as functions of the perturbations in the data $A,\ \ldots,\ D_{21}$ and $\gamma_{\rm opt}$. In the next section we shall derive linear, condition number based bounds for ΔR and ΔS with respect to perturbations in A, B_2, C_2, D_{12}, D_{21} and $\gamma_{\rm opt}$.

III. Linear Sensitivity Analysis

The essence of our approach is to perform sensitivity analysis of LMI (3) and (4) in a similar way as for proper matrix equations after introducing suitable right hand sides which are slightly perturbed.

Consider first LMI (4). Its structure allows us to analyze only the perturbed inequality

$$\begin{aligned}&(\mathcal{N}_{21}+\Delta\mathcal{N}_{21})^{\top}\\ &\times\left\{\left[\begin{array}{cc}(A+\Delta A)^{\top}(S+\Delta S)+(S+\Delta S)(A+\Delta A) & 0\\ B_1{}^{\top}(S+\Delta S) & 0\end{array}\right]\right.\\ &\left.+\left[\begin{array}{cc}0 & (S+\Delta S)B_1\\ 0 & -\gamma I-\Delta\gamma I\end{array}\right]\right\}\\ &\times(\mathcal{N}_{21}+\Delta\mathcal{N}_{21}):=\bar{\mathcal{P}}^*+\Delta\bar{\mathcal{P}}_1<0\end{aligned} \tag{6}$$

where the matrix $\bar{\mathcal{P}}^*$ is obtained using the nominal LMI

$$\mathcal{N}_{21}{}^{\top}\left[\begin{array}{cc}A^{\top}S^*+S^*A & S^*B_1\\ B_1{}^{\top}S^* & -\gamma_{\rm opt}I\end{array}\right]\mathcal{N}_{21}:=\bar{\mathcal{P}}^*<0 \tag{7}$$

and $\Delta\bar{\mathcal{P}}_1$ is due to the data and closed-loop performance perturbations, the rounding errors and the sensitivity of the interior point method that is used to solve the LMIs.

Within first order terms the perturbed relation (6) may be written as

$$\mathcal{N}_{21}{}^{\top}\overline{\mathcal{W}}\mathcal{N}_{21}+\mathcal{N}_{21}{}^{\top}\overline{\mathcal{W}}\Delta\mathcal{N}_{21}+\Delta\mathcal{N}_{21}{}^{\top}\overline{\mathcal{W}}\mathcal{N}_{21}+\Delta\mathcal{N}_{21}{}^{\top}\overline{\mathcal{W}}\Delta\mathcal{N}_{21} \tag{8}$$

where

$$\begin{aligned}&\overline{\mathcal{W}}=\\ &\left[\begin{array}{cc}A^{\top}S^*+S^*A+A^{\top}\Delta S+\Delta SA+\Delta A^{\top}S^*+S^*\Delta A & 0\\ B_1{}^{\top}S^*+B_1\Delta S & 0\end{array}\right]\\ &+\left[\begin{array}{cc}0 & S^*B_1+\Delta SB_1\\ 0 & -\gamma_{\rm opt}I-\Delta\gamma_{\rm opt}I\end{array}\right].\end{aligned}$$

Using relation (19) one has

$$\begin{aligned}\Delta\bar{\mathcal{P}}_1 &= \mathcal{N}_{21}{}^{\top}\Upsilon_S\mathcal{N}_{21}\mathcal{N}_{21}{}^{\top}(\mathcal{P}^*+\Upsilon_S)\Delta\mathcal{N}_{21}+\\ &\quad+\Delta\mathcal{N}_{21}{}^{\top}(\mathcal{P}^*+\Upsilon_S)\mathcal{N}_{21}+\Psi_S)\Delta\mathcal{N}_{21} &(9)\\ &\quad+\Delta\mathcal{N}_{21}{}^{\top}(\mathcal{P}^*+\Upsilon_S)\Delta\mathcal{N}_{21} &(10)\end{aligned}$$

where $\bar{\mathcal{P}}^*=\mathcal{N}_{21}{}^{\top}\mathcal{P}^*\mathcal{N}_{21},\ \Upsilon_S=\Delta_S+\Omega_S$,

$$\Delta_S=\left[\begin{array}{cc}A^{\top}\Delta S+\Delta SA & \Delta SB_1\\ B_1{}^{\top}\Delta S & 0\end{array}\right]$$

$$\Omega_S=\left[\begin{array}{cc}\Delta A^{\top}S^*+S^*\Delta A & 0\\ 0 & -\Delta\gamma_{\rm opt}I\end{array}\right].$$

Neglecting the second and higher order terms in (9) one obtains

$$\begin{aligned}\Delta\bar{\mathcal{P}}_1 &= \mathcal{N}_{21}{}^{\top}\Delta_S\mathcal{N}_{21}+\mathcal{N}_{21}{}^{\top}\Omega_S\mathcal{N}_{21} &(11)\\ &\quad+\Delta\mathcal{N}_{21}{}^{\top}\mathcal{P}^*\mathcal{N}_{21}+\mathcal{N}_{21}{}^{\top}\mathcal{P}^*\Delta\mathcal{N}_{21}.\end{aligned}$$

Setting $\mathcal{P}^*\mathcal{N}_{21}=\tilde{\mathcal{N}}_{21}$ and $\mathcal{N}_{21}{}^{\top}\mathcal{P}^*=\tilde{\mathcal{N}}_{21}^*$ it follows that

$$\begin{aligned}&\mathrm{vec}(\Delta\mathcal{N}_{21}{}^{\top}\tilde{\mathcal{N}}_{21}+\tilde{\mathcal{N}}_{21}^*\Delta\mathcal{N}_{21})= &(12)\\ &[(\tilde{\mathcal{N}}_{21}{}^{\top}\otimes I)\Pi_{(n+l),n^2}+(I\otimes\tilde{\mathcal{N}}_{21}^*)]\mathrm{vec}(\Delta\mathcal{N}_{21}).\end{aligned}$$

Relation (11) may be written in a vector form as

$$\begin{aligned}&(\mathcal{N}_{21}{}^{\top}\otimes\mathcal{N}_{21}{}^{\top})\mathrm{vec}(\Delta_S)+(\mathcal{N}_{21}{}^{\top}\otimes\mathcal{N}_{21}{}^{\top})\mathrm{vec}(\Omega_S) &(13)\\ &+\mathcal{N}_{S\Omega}\mathrm{vec}(\Delta\mathcal{N}_{21})=\mathrm{vec}(\Delta\bar{\mathcal{P}}_1)\end{aligned}$$

where

$$\mathrm{vec}(\Delta_S)=\left[\begin{array}{c}I\otimes A^{\top}+A^{\top}\otimes I\\ B_1{}^{\top}\otimes I\\ I\otimes B_1{}^{\top}\\ 0\end{array}\right]\mathrm{vec}(\Delta S):=V\Delta s$$

$$\mathrm{vec}(\Omega_S) = \left[\begin{array}{cc} (I\otimes S^*)+(S^*\otimes I)\Pi_{n^2} & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & -e_{l^3} \end{array}\right] \times \left[\begin{array}{c} \mathrm{vec}(\Delta A) \\ \Delta\gamma_{\mathrm{opt}} \end{array}\right] := \left[\begin{array}{cc} V_{\mathrm{t1}} & V_{\mathrm{t2}} \end{array}\right] \left[\begin{array}{c} \mathrm{vec}(\Delta A) \\ \Delta\gamma_{\mathrm{opt}} \end{array}\right] \quad (14)$$

and

$$\mathcal{N}_{S\Omega} = (\tilde{\mathcal{N}}_{21}{}^{\top}\otimes I)\Pi_{(n+l),n^2} + (I\otimes\tilde{\mathcal{N}}_{21}^*).$$

Thus we have

$$V_{\mathrm{s}}\Delta s + V_{\mathrm{ts1}}\mathrm{vec}(\Delta A) + V_{\mathrm{ts2}}\Delta\gamma_{\mathrm{opt}} + \mathcal{N}_{S\Omega}\mathrm{vec}(\Delta\mathcal{N}_{21}) = \mathrm{vec}(\Delta\bar{\mathcal{P}}_1) \quad (15)$$

where

$$V_{\mathrm{s}} = (\mathcal{N}_{21}{}^{\top}\otimes\mathcal{N}_{21}{}^{\top})V, \quad V_{\mathrm{ts1}} = (\mathcal{N}_{21}{}^{\top}\otimes\mathcal{N}_{21}{}^{\top})V_{\mathrm{t1}}$$
$$V_{\mathrm{ts2}} = (\mathcal{N}_{21}{}^{\top}\otimes\mathcal{N}_{21}{}^{\top})V_{\mathrm{t2}}.$$

It is well known [6] that the perturbation bound for the projector $\mathcal{N}_{21}$ may be written as

$$\|\Delta\mathcal{N}_{21}\|_2 \le \|[C_2,\ D_{21}]^{\dagger}\|_2\|[\Delta C_2,\ \Delta D_{21}]\|_2. \quad (16)$$

Using the fact that $\|\mathrm{vec}(M)\|_2 = \|M\|_{\mathcal{F}}$, we finally obtain the relative perturbation bound for S^*

$$\frac{\|\Delta S\|_{\mathcal{F}}}{\|S^*\|_{\mathcal{F}}} \le \frac{1}{\|S^*\|_{\mathcal{F}}}\left(V_{\mathrm{ab1}}\frac{\|\Delta A\|_{\mathcal{F}}}{\|A\|_{\mathcal{F}}} + V_{\mathrm{ab2}}\frac{|\Delta\gamma_{\mathrm{opt}}|}{|\gamma_{\mathrm{opt}}|}\right) + \frac{1}{\|S^*\|_{\mathcal{F}}}\left(V_{\mathrm{cd}}\frac{\|[\Delta C_2,\ \Delta D_{21}]\|_{\mathcal{F}}}{\|[C_2,\ D_{21}]\|_{\mathcal{F}}} + V_1\frac{\|\Delta\bar{\mathcal{P}}_1\|_{\mathcal{F}}}{\|\bar{\mathcal{P}}^*\|_{\mathcal{F}}}\right) \quad (17)$$

where

$$\frac{V_{\mathrm{ab1}}}{\|S^*\|_{\mathcal{F}}} = \frac{\|V_{\mathrm{s}}^{\dagger}\|_2\|V_{\mathrm{ts1}}\|_2\|A\|_{\mathcal{F}}}{\|S^*\|_{\mathcal{F}}}$$
$$\frac{V_{\mathrm{ab2}}}{\|S^*\|_{\mathcal{F}}} = \frac{\|V_{\mathrm{s}}^{\dagger}\|_2\|V_{\mathrm{ts2}}\|_2|\gamma_{\mathrm{opt}}|}{\|S^*\|_{\mathcal{F}}}, \quad \frac{V_1}{\|S^*\|_{\mathcal{F}}} = \frac{\|V_{\mathrm{s}}^{\dagger}\|_2\|\bar{\mathcal{P}}^*\|_{\mathcal{F}}}{\|S^*\|_{\mathcal{F}}}$$
$$\frac{V_{\mathrm{cd}}}{\|S^*\|_{\mathcal{F}}} = \frac{\|V_{\mathrm{s}}^{\dagger}\|_2\|\mathcal{N}_{S\Omega}\|_2\|[C_2,\ D_{21}]^{\dagger}\|_{\mathcal{F}}\|[C_2,\ D_{21}]\|_{\mathcal{F}}}{\|S^*\|_{\mathcal{F}}}$$

are the relative condition numbers of LMI (4) with respect to the perturbations in the data.

In a similar way we can obtain a relative perturbation bound for the solution R^* of the LMI (3). In this case we consider the perturbed inequality

$$(\mathcal{N}_{12}+\Delta\mathcal{N}_{12})^{\top} \times\left\{\left[\begin{array}{cc} (A+\Delta A)(R^*+\Delta R)+(R^*+\Delta R)(A+\Delta A)^{\top} & 0 \\ C_1(R^*+\Delta R) & 0 \end{array}\right] + \left[\begin{array}{cc} 0 & (R^*+\Delta R)C_1 \\ 0 & -\gamma_{\mathrm{opt}}I-\Delta\gamma_{\mathrm{opt}}I \end{array}\right]\right\} \times(\mathcal{N}_{12}+\Delta\mathcal{N}_{12}) := \bar{\mathcal{Q}}^* + \Delta\bar{\mathcal{Q}}_1 < 0 \quad (18)$$

where

$$\mathcal{N}_{12}{}^{\top}\left[\begin{array}{cc} AR^*+R^*A^{\top} & R^*C_1{}^{\top} \\ C_1R^* & -\gamma_{\mathrm{opt}}I \end{array}\right]\mathcal{N}_{12} := \bar{\mathcal{Q}}^* < 0. \quad (19)$$

Here, instead of Δ_S and Ω_S we have

$$\Delta_R = \left[\begin{array}{cc} A\Delta R+\Delta RA^{\top} & \Delta RC_1{}^{\top} \\ C_1\Delta R & 0 \end{array}\right]$$
$$\Omega_R = \left[\begin{array}{cc} \Delta AR^*+R^*\Delta A^{\top} & 0 \\ 0 & -\Delta\gamma_{\mathrm{opt}}I \end{array}\right]$$

and thus

$$\mathrm{vec}(\Delta_R) = \left[\begin{array}{c} I\otimes A + A\otimes I \\ C_1\otimes I \\ I\otimes C_1 \\ 0 \end{array}\right]\mathrm{vec}(\Delta R) := T\Delta r$$

$$\mathrm{vec}(\Omega_R) = \left[\begin{array}{cc} (R^*\otimes I)+(I\otimes R^*)\Pi_{n^2} & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & -e_{p^3} \end{array}\right] \times \left[\begin{array}{c} \mathrm{vec}(\Delta A) \\ \Delta\gamma_{\mathrm{opt}} \end{array}\right] = \left[\begin{array}{cc} T_{\mathrm{t1}} & T_{\mathrm{t2}} \end{array}\right] \left[\begin{array}{c} \mathrm{vec}(\Delta A) \\ \Delta\gamma_{\mathrm{opt}} \end{array}\right].$$

Denote

$$\bar{\mathcal{Q}}^* = \mathcal{N}_{12}{}^{\top}\mathcal{Q}^*\mathcal{N}_{12},\ \mathcal{Q}^*\mathcal{N}_{12} = \tilde{\mathcal{N}}_{12},\ \mathcal{N}_{12}{}^{\top}\mathcal{Q}^* = \tilde{\mathcal{N}}_{12}^*$$
$$\mathcal{N}_{R\Omega} = (\tilde{\mathcal{N}}_{12}{}^{\top}\otimes I)\Pi_{(n+p),n^2} + (I\otimes\tilde{\mathcal{N}}_{12}^*)$$
$$T_{\mathrm{r}} = (\mathcal{N}_{12}{}^{\top}\otimes\mathcal{N}_{12}{}^{\top})T,\ T_{\mathrm{tr1}} = (\mathcal{N}_{12}{}^{\top}\otimes\mathcal{N}_{12}{}^{\top})T_{\mathrm{t1}}$$
$$T_{\mathrm{tr2}} = (\mathcal{N}_{12}{}^{\top}\otimes\mathcal{N}_{12}{}^{\top})T_{\mathrm{t2}}.$$

Having in mind that

$$\|\Delta\mathcal{N}_{12}\|_{\mathcal{F}} \le \|[B_2{}^{\top},\ D_{12}{}^{\top}]^{\dagger}\|_{\mathcal{F}}\|[\Delta B_2{}^{\top},\ \Delta D_{12}{}^{\top}]\|_{\mathcal{F}}$$

we obtain the relative perturbation bound for R^*

$$\frac{\|\Delta R\|_{\mathcal{F}}}{\|R^*\|_{\mathcal{F}}} \le \frac{1}{\|R^*\|_{\mathcal{F}}}\left(T_{\mathrm{ac1}}\frac{\|\Delta A\|_{\mathcal{F}}}{\|A\|_{\mathcal{F}}} + T_{\mathrm{ac2}}\frac{|\Delta\gamma_{\mathrm{opt}}|}{|\gamma_{\mathrm{opt}}|}\right) + \frac{1}{\|R^*\|_{\mathcal{F}}}\left(T_{\mathrm{bd}}\frac{\|[\Delta B_2{}^{\top},\ \Delta D_{12}{}^{\top}]\|_{\mathcal{F}}}{\|[B_2{}^{\top},\ D_{12}{}^{\top}]\|_{\mathcal{F}}} + T_1\frac{\|\Delta\bar{\mathcal{Q}}_1\|_{\mathcal{F}}}{\|\bar{\mathcal{Q}}^*\|_{\mathcal{F}}}\right) \quad (20)$$

where

$$\frac{T_{\mathrm{ac1}}}{\|R^*\|_{\mathcal{F}}} = \frac{\|T_{\mathrm{r}}^{\dagger}\|_2\|T_{\mathrm{tr1}}\|_2\|A\|_{\mathcal{F}}}{\|R^*\|_{\mathcal{F}}}$$
$$\frac{T_{\mathrm{ac2}}}{\|R^*\|_{\mathcal{F}}} = \frac{\|T_{\mathrm{r}}^{\dagger}\|_2\|T_{\mathrm{tr2}}\|_2|\gamma_{\mathrm{opt}}|}{\|R^*\|_2}, \quad \frac{T_1}{\|R^*\|_{\mathcal{F}}} = \frac{\|T_r^{\dagger}\|_2\|\bar{\mathcal{Q}}^*\|_{\mathcal{F}}}{\|R^*\|_{\mathcal{F}}}$$
$$\frac{T_{\mathrm{bd}}}{\|R^*\|_{\mathcal{F}}} = \frac{\|T_r^{\dagger}\|_2\|\mathcal{N}_{R\Omega}\|_2\|[B_2{}^{\top},\ D_{12}{}^{\top}]^{\dagger}\|_{\mathcal{F}}\|B_2{}^{\top},\ D_{21}{}^{\top}\|_{\mathcal{F}}}{\|R^*\|_{\mathcal{F}}}$$

are the relative condition numbers of LMI (3).

IV. Numerical Example

Consider the continuous-time system (1) with

$$A=\begin{bmatrix} 0 & 1 \\ -k/m & -c/m \end{bmatrix},\ B_1=\begin{bmatrix} 0 & 0 & 0 \\ -pm & -pc/m & -pk/m \end{bmatrix}$$

$$B_2=\begin{bmatrix} 0 \\ 1/m \end{bmatrix},\ C_1=\begin{bmatrix} -k/m & -c/m \\ 0 & c \\ k & 0 \end{bmatrix}$$

$$C_2=\begin{bmatrix} 1 & 0 \end{bmatrix},\ D_{11}=\begin{bmatrix} -pm & -pc/m & -pk/m \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

$$D_{12}=\begin{bmatrix} 1/m \\ 0 \\ 0 \end{bmatrix},\ D_{21}=\begin{bmatrix} 0 & 0 & 0 \end{bmatrix}$$

and $m=3$, $c=1$, $k=2$, $pm=0.4$, $pc=0.2$, $pk=0.3$. The perturbations in the data are chosen as

$$\Delta A=A\times 10^{-i},\ \Delta B_1=B_1\times 10^{-i},\ \Delta B_2=B_2\times 10^{-i}$$

$$\Delta C_1=C_1\times 10^{-i},\ \Delta C_2=C_2\times 10^{-i},\ \Delta D_{11}=D_{11}\times 10^{-i}$$

$$\Delta D_{12}=D_{12}\times 10^{-i},\ \Delta\gamma_{\rm opt}=10^{-i}\times\gamma_{\rm opt}$$

for $i=8,7,\ldots,4$.

The perturbed solutions $R^*+\Delta R$ and $S^*+\Delta S$ are computed using the LMI Control Toolbox of MATLAB [5]. The optimal closed-loop performance obtained is $\gamma_{\rm opt}=0.4191$. The relative perturbations in the solutions R^* and S^* of (3), (4) are estimated using the perturbation bounds (20) and (17), respectively.

The results obtained for different values of i are shown in the following table:

i	$\frac{\|\Delta S\|_{\mathcal{F}}}{\|S^*\|_{\mathcal{F}}}$	Bound(17)	$\frac{\|\Delta R\|_{\mathcal{F}}}{\|R^*\|_{\mathcal{F}}}$	Bound(20)
8	$1.2\ 10^{-7}$	$3.8\ 10^{-7}$	$1.0\ 10^{-7}$	$1.5\ 10^{-7}$
7	$1.7\ 10^{-7}$	$3.8\ 10^{-6}$	$1.9\ 10^{-7}$	$1.5\ 10^{-6}$
6	$4.1\ 10^{-6}$	$3.8\ 10^{-5}$	$8.2\ 10^{-6}$	$1.5\ 10^{-5}$
5	$1.9\ 10^{-5}$	$3.8\ 10^{-4}$	$9.9\ 10^{-5}$	$1.5\ 10^{-4}$
4	$2.0\ 10^{-4}$	$3.8\ 10^{-3}$	$1.0\ 10^{-4}$	$1.5\ 10^{-3}$

V. Conclusions

Linear sensitivity analysis of the LMI arising in the continuous-time $\mathcal{H}_\infty$ control problem is done. Condition number based perturbation bounds are obtained in a similar way as for matrix equations, introducing a slightly perturbed right hand side in LMI. A numerical example is presented illustrating the accuracy of the proposed LMI perturbation bounds.

References

[1] S. Boyd, L. El Ghaoui, F. Feron, and V. Balakrishnan. *Linear Matrix Inequalities in Systems and Control Theory*, SIAM, Philladelphia, 1992.

[2] N.D. Christov, M.M. Konstantinov, and P.Hr. Petkov. Improved perturbation bounds for the continuous-time H_∞-optimization problem. *Proc. 16th IFAC World Congress*, Prague, 4-8 July 2005, paper Mo-A01-TP/6.

[3] J.C. Doyle, K. Glover, P.P. Khargonekar, and B.A. Francis. State-Space Solutions to Standard H_2 and H_∞ Control Problems. *IEEE Transactions on Automatic Control* **34** (1989) 831–847.

[4] P. Gahinet and P. Apkarian, A Linear Matrix Inequality Approach to H_∞ Control, *Int. J. Robust Non. Contr.* **4** (1994) 421–448.

[5] P. Gahinet, A. Nemirovski, A. Laub, and M. Chilali. *LMI Control Toolbox for Use with MATLAB*, MathWorks, 1992.

[6] G. Steward and J. G. Sun. *Matrix Perturbation Theory*, Academic Press, N.Y., 1990.

[7] Y. Nesterov and A. Nemirovski. *Interior-Point Polynomial Algorithms in Convex Programming*, SIAM, Philladelphia, 1994.

[8] I. R. Peterson, B. D. O Anderson, and E. A. Jonkheere. A first principles solution to the non-singular H_∞ control problem, *Int. J. Robust Non. Contr.* **1** (1991) 171–185.

[9] K. Zhou, J.C. Doyle, and K. Glover. *Robust and Optimal Control*, Prentice-Hall, Upper Saddle River, N.J., 1995.

4 Revisited H-inf control with transients: LMI based time-invariant output-feedback controllers

Revisited H_∞ control with transients: LMI based time-invariant output-feedback controllers

Dmitry V. Balandin
Department of Numerical and Functional Analysis
Nizhny Novgorod State University
Gagarin ave., 23, Nizhny Novgorod, 603950 Russia
Email: dbalandin@yandex.ru

Mark M. Kogan
Department of Mathematics
Architecture and Civil Engineering University
Il'yinskaya str., 65, Nizhny Novgorod, 603950 Russia
Email: mkogan@nngasu.ru

***Abstract*—Attenuating both exogenous signals and initial disturbances caused by unknown initial conditions is considered in the framework of a so called problem of H_∞ control with transients. Applying an LMI approach, instead of the Riccati equations one, in characterizing the performance measure that is the worst-case norm of the regulated output over all exogenous signals and initial states allows one to synthesize a time-invariant, instead of the time-varying, output-feedback controller for which the performance measure of the closed-loop system is less than a prescribed number. State-space formulae for all time-invariant state- and output-feedback controllers in the problem of the H_∞ control with transients are also presented.**

***Index Terms*—H_∞ control, unknown initial conditions, linear matrix inequality, output-feedback controller**

I. Introduction

The classical H_∞ control theory [1], [2] defines the control law for which the performance measure that is the worst-case norm of the regulated output over all exogenous signals less than a prescribed number. It is usually assumed that the plant initial state is zero. However, there exist situations when the plant initial state is possibly nonzero and unknown. The nonzero initial state causes an additional unknown disturbance. In [3], [4], γ-optimal control which minimizes the worst-case norm of the regulated output over all initial states in the disturbance-free system was considered. When two above reasons are available, it is worth-while to synthesize a control law that would provide attenuating both exogenous and initial disturbances.

Reference [5] introduced a performance measure that is the induced norm of the regulated output over all exogenous signals and initial states for finite and infinite horizons (see also [6], [7]). The performance measure is parameterized by a weighting matrix R reflecting the relative importance of the uncertainty in the initial state contrary to the uncertainty in the exogenous signal. The problem is to synthesize controllers for which the performance measure of the closed-loop system is less than a prescribed number. Since the authors of [5] from the very outset restrict their attention to the central solution to the corresponding H_∞-like problem based on the Riccati equations and since the central controller for the problem under consideration turned out, even in the infinite horizon case, to be a linear time-varying output-feedback controller, they obtained necessary and sufficient conditions for the existence of a linear time-varying output-feedback controller. This observer-based controller is of the form

$$\begin{aligned}\frac{d\hat{x}}{dt} = (A+\gamma^{-2}B_1B_1^TP)\hat{x} + [I-\gamma^{-2}Q(t)P]^{-1}\times \\ Q(t)C_2^T(y-C_2\hat{x}) + B_2u, \\ u = -B_2^TP\hat{x},\end{aligned} \tag{1}$$

where P is the stabilizing solution to H_∞-type algebraic Riccati equation such that $P < \gamma^2 R$, $Q(t)$ is the solution to H_∞-type differential Riccati equation for the finite time horizon with initial condition $Q(0) = R^{-1}$ such that the unforced linear time-varying system

$$\dot{p} = [A - Q(t)(C_2^TC_2 - \gamma^{-2}C_1^TC_1)]p$$

is exponentially stable, and finally "spectral radius" function $\{1-\gamma^{-2}\rho[Q(t)P]\}^{-1} > 0$ for all $t \geq 0$ and is bounded. From the numerical point of view, constructing such a controller based on solving these algebraic and differential Riccati equations coupled to the nonstationary algebraic inequality seems to be a rather complicated problem, and finding an optimal controller is very problematical. Quantitative results concerning another time-varying output-feedback controller for a slightly different H_∞ control problem with initial conditions were obtained in [8].

At the same time, there exists a whole set of controllers that yield the prescribed performance. Among these controllers there maybe time-invariant output-feedback ones. The present paper just shows that this is the case. Necessary and sufficient conditions for the existence of time-invariant state- and output-feedback controllers in the H_∞ control problem with transients are derived in the terms of LMIs, and state-space formulae for all such controllers are presented. The role of the weighting matrix is revealed in the trade-off between H_∞- and γ-optimal controls. More precisely, necessary and sufficient conditions in the form of a fundamental inequality for the weighting matrix is obtained under which the above trade-off takes place.

II. Performance measure

Let the asymptotically stable system be described by the equations

$$\begin{aligned} \dot{x} &= Ax + Bv, \quad x(0) = x_0, \\ z &= Cx + Dv, \end{aligned} \tag{2}$$

where $x \in \mathcal{R}^{n_x}$ is the state, $v \in \mathcal{R}^{n_v}$ is the exogenous input, $z \in \mathcal{R}^{n_z}$ is the regulated output. It is assumed that the exogenous disturbance $v = v(t) \in L_2[0, \infty)$ and that the plant initial state x_0 is unknown. The performance measure is defined as the worst-case norm of the regulated output over all admissible exogenous signals and initial states [5], i.e.,

$$\gamma_w = \sup_{\|v\|^2 + x_0^T R x_0 \neq 0} \frac{\|z\|}{(\|v\|^2 + x_0^T R x_0)^{1/2}}, \tag{3}$$

where $\|\cdot\|$ denotes L_2-norm of the corresponding function and $R = R^T > 0$ is a given weighting matrix.

If the initial state is zero, γ_w is equal to the worst-case norm of the regulated output over all admissible exogenous signals, i.e., $\gamma_w = \gamma_\infty$, where

$$\gamma_\infty = \sup_{\|v\| \neq 0} \frac{\|z\|}{\|v\|} = \sup_{\omega \in (-\infty, \infty)} \|H(j\,\omega)\| = \|H\|_\infty,$$

$H(s) = D + C(sI - A)^{-1}B$ is the transfer function matrix, $j = \sqrt{-1}$, $\|H\| = \max\limits_i \sigma_i(H)$, σ_i is the ith singular value of the matrix H, and $\|\cdot\|_\infty$ is the ∞-norm in the space of $H(s)$ such that $\sup_{\mathrm{Re}\, s \geq 0} \|H(s)\| < \infty$.

If the exogenous input is zero, γ_w is the worst-case norm of the regulated output over all admissible initial states, i.e., $\gamma_w = \gamma_0(R)$, where

$$\gamma_0(R) = \sup_{x_0 \neq 0} \frac{\|z\|}{(x_0^T R x_0)^{1/2}}. \tag{4}$$

It is shown in [5] that the performance measure of the system (2) with $D = 0$ satisfies inequality $\gamma_w < \gamma$ for a given γ if and only if there exists a symmetric matrix P such that

$$A^T P + PA + \gamma^{-2} PBB^T P + C^T C = 0,$$

$A + \gamma^{-2} BB^T P$ is asymptotically stable, and $P < \gamma^2 R$. Even though in the characterization of the standard H_∞ norm the Riccati equation can be substituted for the corresponding Riccati inequality, that is equivalent, by Schur lemma, to LMI, it does not yet follow from this that $\gamma_w < \gamma$ is equivalent to these two LMIs. This fact requires a separate proof.

Theorem 1: The performance measure of the system (2) is less than a prescribed value γ if and only if there exists a $(n_x \times n_x)$-matrix $X = X^T > 0$ such that LMIs

$$\begin{pmatrix} A^T X + XA & XB & C^T \\ B^T X & -\gamma^2 I & D^T \\ C & D & -I \end{pmatrix} < 0, \quad X < \gamma^2 R \tag{5}$$

are feasible.

The proof of Theorem 1 utilizes some ideas of the proof of Theorem 1.3 in [5] and is given in Appendix. Since γ_w can equivalently be expressed as

$$\gamma_w = \inf_\gamma \{\gamma : \frac{\|z\|}{(\|v\|^2 + x_0^T R x_0)^{1/2}} < \gamma, \\ \forall\, v \in L_2, \forall\, x_0 \in \mathcal{R}^{n_x}, \|v\|^2 + x_0^T R x_0 \neq 0\},$$

from Theorem 1 it follows that γ_w can be computed as the minimal value of γ for which LMIs (5) are feasible in variables $X = X^T > 0$ and $\gamma^2 > 0$. This is a standard procedure in Matlab which results in a value $\gamma_\varepsilon = \gamma_w + \varepsilon$, where $\varepsilon > 0$ is determined by the accuracy of LMI solvers.

Now we study the performance measure as a function of the weighting matrix R, i.e., $\gamma_w = \gamma_w(R)$. From (3) it follows that if $R_1 \leq R_2$, then $\gamma_w(R_1) \geq \gamma_w(R_2)$. This property was also mentioned in [5]. Further, since

$$\sup_{\|v\|^2 + x_0^T R x_0 \leq 1} \|z\| \geq \sup_{\|v\| = 1,\, x_0 = 0} \|z\| = \gamma_\infty, \\ \sup_{\|v\|^2 + x_0^T R x_0 \leq 1} \|z\| \geq \sup_{x_0^T R x_0 = 1,\, v \equiv 0} \|z\| = \gamma_0(R),$$

we arrive at $\gamma_w(R) \geq \max\{\gamma_\infty, \gamma_0(R)\}$. The next property of the performance measure plays a special role and is proven in Appendix.

Theorem 2: Let the system and the performance measure be as above. Then $\gamma_w(R) = \gamma_\infty$ if and only if $\lambda_{max}(R^{-1} R_*) \leq 1$, where $R_* = (1/\gamma_\infty^2) \lim\limits_{\varepsilon \to +0} X(\varepsilon)$ and $X(\varepsilon)$ is a stabilizing solution of the Riccati equation

$$A^T X + XA + C^T C + \varepsilon^2 I + \\ (C^T D + XB)[(\gamma_\infty + \varepsilon)^2 I - D^T D]^{-1} (C^T D + XB)^T = 0. \tag{6}$$

Corollary 1: The performance measure (3) for the system (2) is the trade-off between $\|H\|_\infty$ and $\gamma_0(R)$ if and only if the following inequality holds

$$\lambda_{max}(R^{-1} R_*) > 1, \tag{7}$$

where R_* is defined in Theorem 2.

To compute the matrix R_* it is sufficient to find both the minimal $\gamma_* \approx \gamma_\infty$ and the corresponding matrix X_* satisfying the first inequality in (5), and then $R_* \approx X_*/\gamma_*^2$.

III. State-feedback control

Consider the controlled plant

$$\begin{aligned} \dot{x} &= Ax + B_1 v + B_2 u, \\ z &= C_1 x + D_{11} v + D_{12} u, \end{aligned} \tag{8}$$

where $x \in \mathcal{R}^{n_x}$ is the state, $v \in \mathcal{R}^{n_v}$ is the exogenous input, $u \in \mathcal{R}^{n_u}$ is the control input, and $z \in \mathcal{R}^{n_z}$ is the regulated output. The problem is to find a linear state-feedback control

$$u = \Theta x \tag{9}$$

for which the performance measure $\gamma_w(\Theta)$ of the closed-loop system is less than a given $\gamma > 0$.

Substituting the matrix of the closed-loop system into (5), multiplying the first inequality by $\mathrm{diag}\,(X^{-1}, I, I)$ from the left and from the right, introducing the new variables $Y =$

X^{-1} and $Z = \Theta Y$, and changing the second inequality by using Schur lemma, we arrive at the following statement.

Theorem 3: There exists an admissible state-feedback controller such that $\gamma_w(\Theta) < \gamma$ if and only if the LMIs

$$\begin{pmatrix} YA^T + AY + B_2Z + Z^TB_2^T & * & * \\ B_1^T & -\gamma^2 I & * \\ C_1Y + D_{12}Z & D_{11} & -I \end{pmatrix} < 0,$$
$$\begin{pmatrix} Y & I \\ I & \gamma^2 R \end{pmatrix} > 0 \tag{10}$$

are feasible in the variables $Y = Y^T > 0$ and Z, where $*$ stands for notation of the corresponding entry of the symmetric matrix. In this case, gain matrix Θ is computed as $\Theta = ZY^{-1}$, where (Y, Z) is a solution of (10) for the given γ.

Define an optimal state-feedback controller by the inequality $\gamma_w(\Theta) < \gamma_w^*(1+\varepsilon)$ for any sufficiently small $\varepsilon > 0$, where $\gamma_w^* = \inf_{\Theta} \gamma_w(\Theta)$. To numerically compute γ_w^* and the optimal gain matrix Θ_* it is required to find a minimal value of γ and the corresponding Y_*, Z_* for which (10) are feasible with regard to $Y = Y^T > 0$, Z, and $\gamma^2 > 0$ and then compute $\Theta_* = Z_*Y_*^{-1}$.

Remark 1: Note that the gain matrix for the standard H_∞-optimal state-feedback controller can be computed as $\Theta_\infty = Z_*Y_*^{-1}$, where (Y_*, Z_*) is the solution of the first inequality in (10) with the minimal value of $\gamma \approx \bar{\gamma}_\infty$. From the second inequality in (10) it immediately follows that for $R > R_* = \bar{\gamma}_\infty^{-2} Y_*^{-1}$, we get $\gamma_w^*(R) = \bar{\gamma}_\infty$ and, hence, H_∞-optimal state-feedback controller with transients coincides with the standard H_∞-optimal controller. Also note that γ-optimal state-feedback controller (see [3]) is computed as $\Theta_0(R) = Z_0Y_0^{-1}$, where (Y_0, Z_0) is the solution of LMIs (10) in the first of which the second row and column are deleted, with the minimal value of $\gamma \approx \bar{\gamma}_0(R)$.

IV. Output-feedback control

Let the plant be described by the equations

$$\begin{aligned} \dot{x} &= Ax + B_1v + B_2u, \\ z &= C_1x + D_{11}v + D_{12}u, \\ y &= C_2x + D_{21}v, \end{aligned} \tag{11}$$

where $x \in R^{n_x}$ is the plant state, $v \in \mathcal{R}^{n_v}$ is the exogenous input, $u \in R^{n_u}$ is the control input, $z \in R^{n_z}$ is the regulated output, $y \in R^{n_y}$ is the measurable output, and the full order dynamic output-feedback controller be described by

$$\begin{aligned} \dot{x}_r &= A_rx_r + B_ry\ , \quad x_r(0) = 0\ , \\ u &= C_rx_r + D_ry\ , \end{aligned} \tag{12}$$

where $x_r \in R^{n_x}$ is the controller state. Denote the gain matrix of the controller as

$$\Theta = \begin{pmatrix} A_r & B_r \\ C_r & D_r \end{pmatrix}.$$

The problem is to synthesize a time-invariant output-feedback controller such that the performance measure of the closed-loop system satisfies the inequality $\gamma_w(\Theta) < \gamma$ with a given $\gamma > 0$.

Theorem 4: There exists an admissible output-feedback controller such that $\gamma_w(\Theta) < \gamma$ if and only if the LMIs

$$N_1^T \begin{pmatrix} A^TX_{11} + X_{11}A & * & * \\ B_1^TX_{11} & -\gamma^2 I & * \\ C_1 & D_{11} & -I \end{pmatrix} N_1 < 0,$$
$$N_2^T \begin{pmatrix} Y_{11}A^T + AY_{11} & * & * \\ C_1Y_{11} & -I & * \\ B_1^T & D_{11}^T & -\gamma^2 I \end{pmatrix} N_2 < 0, \tag{13}$$
$$\begin{pmatrix} X_{11} & I \\ I & Y_{11} \end{pmatrix} \geq 0, \quad X_{11} < \gamma^2 R$$

are feasible in $(n_x \times n_x)$-matrices $X_{11} = X_{11}^T > 0$ and $Y_{11} = Y_{11}^T > 0$, where columns of the matrices N_1 and N_2 form bases of kernels of matrices $(C_2\ D_{21}\ 0)$ and $(B_2^T\ D_{12}^T\ 0)$, respectively.

The proof of Theorem 4 is based on Theorem 1 and manipulations using LMI technique and given in Appendix. The procedure of computing gain matrix Θ is following: find matrices X_{11} and Y_{11}, construct matrix X using formula (23), and solve LMI (18) with respect to Θ. Note that the problem is rendered tractable under the assumption that the initial conditions of the controller states are zero, otherwise the problem is not reduced to convex optimization (see [3]).

An optimal output-feedback controller is defined by the inequality $\gamma_w(\Theta) < \gamma_w^*(1+\varepsilon)$ for any sufficiently small $\varepsilon > 0$, where $\gamma_w^* = \inf_{\Theta} \gamma_w(\Theta)$. Analogously to the state-feedback case for $R > R_* = \hat{\gamma}_\infty^{-2} X_{11}^*$, where X_{11}^* is the solution to inequalities in (13), except the last one, with the minimal value of $\gamma \approx \hat{\gamma}_\infty$, we have $\gamma_w^* = \hat{\gamma}_\infty$ and, hence, for such weighting matrices, the optimal output-feedback controller coincides with the standard H_∞-optimal output-feedback controller. This very steady-state central H_∞ controller was derived in [5] (see Corollary 2.5) as a solution to the problem of H_∞ control with transients for sufficiently large R.

V. Illustrative example

Consider a controlled linear oscillator described by the equations (8) with matrices

$$A = \begin{pmatrix} 0 & 1 \\ -1 & -0.1 \end{pmatrix},\ B_1 = B_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix},$$
$$C_1 = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix},\ D_{11} = \begin{pmatrix} 0 \\ 0 \end{pmatrix},\ D_{12} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}.$$

Calculate the standard H_∞-optimal state-feedback controller to find the gain matrix $\Theta_\infty = -(0.5012\ 10.0217)$ and $\bar{\gamma}_\infty = 0.9950$. For γ-optimal controller we find $\Theta_0(\rho) = -(0.4142\ 0.8156)$ and $\bar{\gamma}_0(\rho) = 1.2087/\rho$. Note that in this case the gain matrix $\Theta_0(\rho)$ does not depend on ρ. Fig.1 shows three curves: curve 1 is the performance measure γ_w for the the closed-loop system with Θ_∞; curve 2 is that for the closed-loop system with Θ_0; curve 3 is the optimal performance measure $\gamma_w^*(\rho)$ corresponding to H_∞-optimal state-feedback controller with transients. From this figure it follows that the

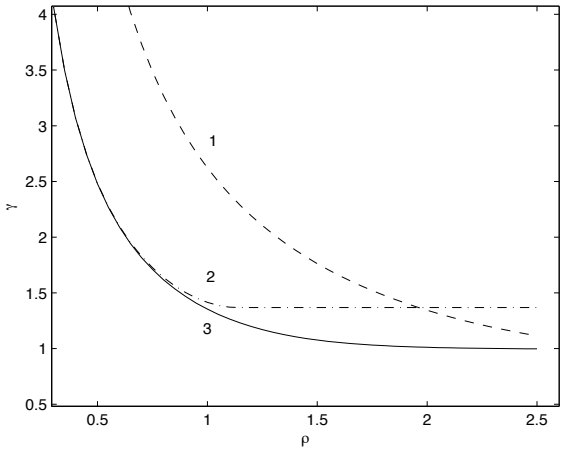


Fig. 1. Performance measure of the closed-loop system under (1) H_∞-optimal, (2) γ-optimal and (3) H_∞-optimal with transients controllers versus parameter ρ in the state-feedback case.

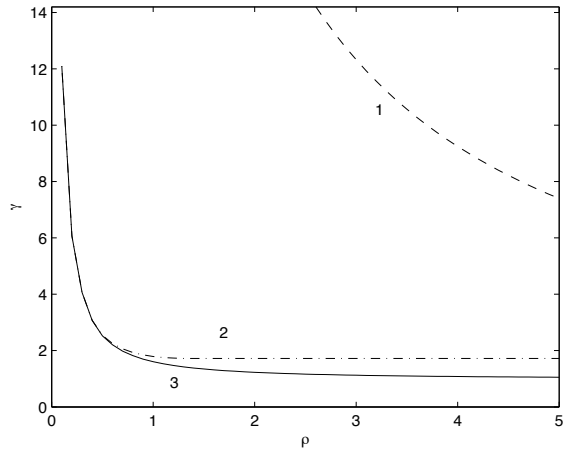


Fig. 2. Performance measure of the closed-loop system under (1) H_∞-optimal, (2) γ-optimal and (3) H_∞-optimal with transients controllers versus parameter ρ in the output-feedback case.

performance measure $\gamma_w^*(\rho)$ is closed to $\bar{\gamma}_0(\rho)$ for small ρ and is closed to $\bar{\gamma}_\infty$ for large ρ, while in a certain range of ρ it is considerably less than both of these.

In the output-feedback case, we obtained the following parameters for H_∞-optimal and γ-optimal controllers

$$\Theta_\infty = \begin{pmatrix} -26,35 & 1,00 & -26,35 \\ -75,84 & -0,0004 & -74,85 \\ -252,4 & 10,00 & -252,9 \end{pmatrix},$$

$$\Theta_0 = \begin{pmatrix} -1,5056 & 0,0334 & 0,6472 \\ 0,0506 & -1,5251 & -1,0082 \\ 0,8990 & -1,3760 & -0,9378 \end{pmatrix}.$$

Fig.2 shows three curves: curve 1 is the performance measure $\gamma_w(\rho)$ for the the closed-loop system with Θ_∞; curve 2 is that for the closed-loop system with Θ_0; curve 3 is the optimal performance measure $\gamma_w^*(\rho)$ corresponding to H_∞-optimal output-feedback controller with transients. Let us compare the values of the performance measure, for example, at $\rho = 4,5$: $\gamma_p(\Theta_\infty) = 8,2203$, $\gamma_p(\Theta_0) = 1,7221$ and $\gamma_p^* = 1,0651$ under H_∞-optimal output-feedback controller with transients for which

$$\Theta_* = \begin{pmatrix} -5,425 & 0,9584 & -5,587 \\ 0,1695 & -0,5665 & 0,4048 \\ -10,93 & 2,142 & -11,42 \end{pmatrix}.$$

Thus, the performance measure under the optimal controller constructed is considerably less then that under the standard H_∞-optimal controller.

VI. Conclusion

This paper presents an LMI approach to H_∞ control incorporating unknown initial conditions. The main contribution of the paper is the new LMI based necessary and sufficient conditions for the existence of a time-invariant output-feedback controller solving the problem of H_∞ control with transients and a procedure of synthesizing optimal controllers. It is also shown that H_∞-optimal control with transients is actually a trade-off between H_∞-control, being optimal under unknown exogenous disturbances and zero initial state, and γ-control, being optimal under zero exogenous signal and unknown initial conditions, if and only if the weighting matrix satisfies a fundamental inequality. If this inequality fails, the performance measure coincides with the H_∞-norm and the trade-off gets broken.

Acknowledgment

This work was supported in part by the Russian Foundation for Basic Research under Grants 07-01-00481, 08-01-00422, 08-01-97034-r.

Appendix A
Proof of Theorem 1

By Schur lemma the first inequality in (5) implies that for all x, v $(|x|^2 + |v|^2 \neq 0)$

$$x^T(A^TX + XA)x + 2x^TXBv +$$
$$(Cx + Dv)^T(Cx + Dv) - \gamma^2|v|^2 < 0.$$

This means that the derivative of the function $V(x) = x^TXx$ along a trajectory of the system (2) satisfies inequality

$$\dot{V} + |z|^2 - \gamma^2|v|^2 < 0 \quad \forall\, x, v\,(|x|^2 + |v|^2 \neq 0).$$

Integrating this inequality over infinite horizon, we get $\|z\|^2 - \gamma^2\|v\|^2 < x_0^TXx_0$. Then the second inequality in (5) implies $\gamma_w < \gamma$.

Let now $\gamma_w < \gamma$. Since $\gamma_\infty \leq \gamma_w$, then $\|H\|_\infty < \gamma$ and, hence, it follows from KYP lemma [9], [10] that the first inequality in (5) is feasible and that there exists a stabilizing solution $X_\gamma = X_\gamma^T > 0$ of the Riccati equation

$$A^TX + XA + C^TC +$$
$$(C^TD + XB)(\gamma^2I - D^TD)^{-1}(C^TD + XB)^T = 0$$

such that matrix $A + B(\gamma^2 I - D^T D)^{-1}(C^T D + X_\gamma B)^T$ is asymptotically stable. Note that $\gamma^2 I - D^T D > 0$ due to the first inequality in (5). This means that the derivative of the function $V(x) = x^T X_\gamma x$ along trajectories of the system (2) satisfies

$$\dot{V} + |z|^2 - \gamma^2 |v_*|^2 = 0, \tag{14}$$

where $v_* = (\gamma^2 I - D^T D)^{-1}(C^T D + X_\gamma B)^T x$.

We will show that there exists a solution of the first inequality in (5) satisfying the second inequality. To that end, at first we show that matrix X_γ satisfies the second inequality in (5). Suppose this is not the case, i.e., there exists $x_0 \neq 0$ such that $x_0^T X_\gamma x_0 \geq \gamma^2 x_0^T R x_0$. Choose the initial state $x(0) = x_0$ and integrate (14) to obtain $\|z\|^2 - \gamma^2 \|v_*\|^2 = x_0^T X_\gamma x_0 \geq \gamma^2 x_0^T R x_0$. This implies $\gamma_w \geq \gamma$ which contradicts the assumption. Thus, $X_\gamma < \gamma^2 R$ and, consequently, $\lambda_{max}(R^{-1/2} X_\gamma R^{-1/2}) = \lambda_{max}(R^{-1} X_\gamma) < \gamma^2$, where $\lambda_{max}(\cdot)$ denotes the maximum eigenvalue of the corresponding matrix.

Suppose that any solution X of the Riccati inequality

$$A^T X + XA + C^T C + (C^T D + XB)(\gamma^2 I - D^T D)^{-1}(C^T D + XB)^T < 0$$

that is equivalent to the first inequality in (5) does not satisfy the inequality $X < \gamma^2 R$, i.e., $\lambda_{max}(R^{-1} X) \geq \gamma^2$ holds. Then there exists a consequence $\varepsilon_n \to 0$ as $n \to \infty$ such that, for stabilizing solutions $X(\varepsilon_n) = X^T(\varepsilon_n) > 0$ to the Riccati equations

$$A^T X + XA + C^T C + (C^T D + XB)(\gamma^2 I - D^T D)^{-1}(C^T D + XB)^T + \varepsilon_n^2 I = 0,$$

the inequalities $\lambda_{max}[R^{-1} X(\varepsilon_n)] \geq \gamma^2$ hold. Since $X(\varepsilon_n) \to X_\gamma$ as $n \to \infty$, the limiting matrix possesses the analogous property, which contradicts to $\lambda_{max}(R^{-1} X_\gamma) < \gamma^2$.

Appendix B
Proof of Theorem 2

Let $\gamma_w(R) = \gamma_\infty$ and $X(\varepsilon)$ be a stabilizing solution of (6). By Schur lemma $X(\varepsilon)$ satisfies the first inequality in (5) for $\gamma = \gamma_\infty + \varepsilon$, where $\varepsilon > 0$. It was shown in the proof of Theorem 1 that the inequality $X(\varepsilon) < \gamma_w^2(R) R$ holds. Taking the limit as $\varepsilon \to 0$ we get $\gamma_\infty^2 R_* \leq \gamma_w^2(R) R$ and, hence, $R \geq R_*$ or, equivalently, $\lambda_{max}(R^{-1} R_*) \leq 1$.

Now, let $\lambda_{max}(R^{-1} R_*) \leq 1$. Denote $R(\varepsilon) = X(\varepsilon)/(\gamma_\infty + \varepsilon)^2$. Since $X(\varepsilon) = (\gamma_\infty + \varepsilon)^2 R(\varepsilon) < (\gamma_\infty + \varepsilon)^2 R$ for $R > R(\varepsilon)$, then $X(\varepsilon)$ also satisfies the second inequality in (5) for $\gamma = \gamma_\infty + \varepsilon$ and $R > R(\varepsilon)$. Now, from Theorem 1 it follows $\gamma_w(R) < \gamma_\infty + \varepsilon$ for $R > R(\varepsilon)$. Taking the limit as $\varepsilon \to 0$, we get $\gamma_w(R) \leq \gamma_\infty$ for $R \geq R_*$. In view of $\gamma_w(R) \geq \gamma_\infty$ for any $R > 0$, we arrive at $\gamma_w(R) = \gamma_\infty$ for $R \geq R_*$.

Appendix C
Proof of Theorem 4

The closed-loop system (11), (12) is described by the equations

$$\begin{aligned} \dot{x}_c &= A_c x_c + B_c v, \\ z &= C_c x_c + D_c v, \end{aligned} \tag{15}$$

where $x_c = \mathrm{col}\,(x, x_r)$,

$$\begin{aligned} A_c &= \begin{pmatrix} A + B_2 D_r C_2 & B_2 C_r \\ B_r C_2 & A_r \end{pmatrix}, \\ B_c &= \begin{pmatrix} B_1 + B_2 D_r D_{21} \\ B_r D_{21} \end{pmatrix}, \\ C_c &= \begin{pmatrix} C_1 + D_{12} D_r C_2 & D_{12} C_r \end{pmatrix}, \\ D_c &= D_{11} + D_{12} D_r D_{21}. \end{aligned} \tag{16}$$

Since the initial state of this system is of the form $\mathrm{col}\,(x_0, 0)$, its performance measure is determined by (3) as before. From the proof of Theorem 1 it immediately follows that the LMI characterization of such a performance measure is reduced to

$$\begin{pmatrix} A_c^T X + X A_c & X B_c & C_c^T \\ B_c^T X & -\gamma^2 I & D_c^T \\ C_c & D_c & -I \end{pmatrix} < 0, \quad X_{11} < \gamma^2 R, \tag{17}$$

where X_{11} is the $(n_x \times n_x)$ top left block of the matrix X.

Further, let us present the matrices of the closed-loop system in the form

$$A_c = A_0 + \mathcal{B}\Theta\mathcal{C}\ , \quad B_c = B_0 + \mathcal{B}\Theta\mathcal{D}_{21}\ ,$$

$$C_c = C_0 + \mathcal{D}_{12}\Theta\mathcal{C}\ , \quad D_c = D_{11} + \mathcal{D}_{12}\Theta\mathcal{D}_{21}\ ,$$

where

$$A_0 = \begin{pmatrix} A & 0 \\ 0 & 0 \end{pmatrix}, B_0 = \begin{pmatrix} B_1 \\ 0 \end{pmatrix}, C_0 = \begin{pmatrix} C_1 & 0 \end{pmatrix},$$

$$\mathcal{B} = \begin{pmatrix} 0 & B_2 \\ I & 0 \end{pmatrix}, \quad \mathcal{C} = \begin{pmatrix} 0 & I \\ C_2 & 0 \end{pmatrix},$$

$$\mathcal{D}_{12} = \begin{pmatrix} 0 & D_{12} \end{pmatrix}, \quad \mathcal{D}_{21} = \begin{pmatrix} 0 \\ D_{21} \end{pmatrix}.$$

Insert these expressions into inequality (17) to present it in the form (see also [11], [12], [4])

$$\Psi + P^{\mathrm{T}} \Theta^T Q + Q^{\mathrm{T}} \Theta P < 0, \tag{18}$$

where

$$\Psi = \begin{pmatrix} A_0^T X + X A_0 & X B_0 & C_0^T \\ B_0^T X & -\gamma I & D_{11}^T \\ C_0 & D_{11} & -\gamma I \end{pmatrix}, \tag{19}$$

$$P = \begin{pmatrix} \mathcal{C} & \mathcal{D}_{21} & 0 \end{pmatrix}, \quad Q = \begin{pmatrix} \mathcal{B}^T X & 0 & \mathcal{D}_{12}^T \end{pmatrix}.$$

Then by elimination lemma, inequality (18) holds for some Θ if and only if

$$\begin{aligned} W_P^T \begin{pmatrix} A_0^T X + X A_0 & X B_0 & C_0^T \\ B_0^T X & -\gamma I & D_{11}^T \\ C_0 & D_{11} & -\gamma I \end{pmatrix} W_P < 0, \\ W_Q^T \begin{pmatrix} A_0^T X + X A_0 & X B_0 & C_0^T \\ B_0^T X & -\gamma I & D_{11}^T \\ C_0 & D_{11} & -\gamma I \end{pmatrix} W_Q < 0, \end{aligned} \tag{20}$$

where W_S stands for notation of a matrix whose columns form any base of the null space of the matrix S. Observe that

$$Q = R \begin{pmatrix} X & 0 & 0 \\ 0 & I & 0 \\ 0 & 0 & I \end{pmatrix}, \quad R = \begin{pmatrix} \mathcal{B}^T & 0 & \mathcal{D}_{12}^T \end{pmatrix}.$$

Hence,

$$W_Q = \begin{pmatrix} X^{-1} & 0 & 0 \\ 0 & I & 0 \\ 0 & 0 & I \end{pmatrix} W_R.$$

Thus, inequalities (20) are equivalent to following two LMIs

$$\begin{aligned} W_P^T \begin{pmatrix} A_0^T X + X A_0 & X B_0 & C_0^T \\ B_0^T X & -\gamma I & D_{11}^T \\ C_0 & D_{11} & -\gamma I \end{pmatrix} W_P < 0, \\ W_R^T \begin{pmatrix} Y A_0^T + A_0 Y & B_0 & Y C_0^T \\ B_0^T & -\gamma I & D_{11}^T \\ C_0 Y & D_{11} & -\gamma I \end{pmatrix} W_R < 0, \end{aligned} \tag{21}$$

where $Y = X^{-1}$. Taking into account the block structures of the above matrices and partitioning X and Y as

$$X = \begin{pmatrix} X_{11} & X_{12} \\ X_{12}^{\mathrm{T}} & X_{22} \end{pmatrix}, \quad Y = \begin{pmatrix} Y_{11} & Y_{12} \\ Y_{12}^{\mathrm{T}} & Y_{22} \end{pmatrix},$$

we arrive at the first pair of inequalities in (13).

According to Frobenius formula, $Y = X^{-1}$ implies

$$Y_{11} = (X_{11} - X_{12} X_{22}^{-1} X_{12}^T)^{-1} \tag{22}$$

which shows there exist reciprocal matrices $X > 0$, $Y > 0$ with given blocks $X_{11} = X_{11}^{\mathrm{T}} > 0$, $Y_{11} = Y_{11}^{\mathrm{T}} > 0$ if and only if $X_{11} - Y_{11}^{-1} \geq 0$, i.e. the third inequality in (13) holds. If this inequality is strict, blocks X_{12} and X_{22} of the corresponding matrix X can be chosen, for example, as

$$X_{12} = X_{22} = X_{11} - Y_{11}^{-1} \ . \tag{23}$$

Thus, inequalities (17) are feasible if and only if inequalities (13) are feasible. This completes the proof.

REFERENCES

[1] G. Zames, "Feedback and optimal sensitivity: model reference transformations, multiplicative seminorms, and approximate inverses," *IEEE Trans. Automat. Control*, vol. 26, pp. 301-320, 1981.

[2] J.C. Doyle, K. Glover, P.P. Khargonekar, and B.F. Francis, "State-space solutions to standard H_2 and H_∞ control problems," *IEEE Trans. Automat. Control*, vol. 34, no. 8, pp. 831-847, 1989.

[3] Balandin, D.V. and Kogan, M.M., 2008, LMI based output-feedback controllers: γ-optimal versus linear quadratic, *Proc. of 17th World IFAC Congress*, Seoul, Korea, 9905–9909.

[4] Balandin, D.V. and Kogan, M.M., 2009, Revisited LQ output-feedback control: minimax controller for a set of initial states," *Int. J. of Control*, **82** (11), 2051–2058.

[5] P.P. Khargonekar, K.M. Nagpal, and K.R. Poolla, "H_∞ control with transients," *SIAM J. Control and Optimization*, vol. 29, no. 6, pp. 1373-1393, 1991.

[6] K. Uchida, A. Kojima, and M. Fujita, "H_∞ control attenuating initial-state uncertainties," *Int. J. of Control*, vol. 65, no. 2, pp. 245-252, 1997.

[7] W.W. Lu, G.J. Balas, and E.B. Lee, "A variational approach to H_∞ control with transients," *IEEE Trans. Automat. Control*, vol. 44, no. 10, pp. 1875-1879, 1999.

[8] Y.K. Foo, "H_∞ control with initial conditions," *IEEE Trans. Circuits and Systems*, vol. 53, no. 9, pp. 867-871, 2006.

[9] V.A. Yakubovich, "Solution of certain matrix inequalities in the stability theory of nonlinear control systems," *Soviet Math. Dokl.*, vol. 3, pp. 620-623, 1962.

[10] M. Green, and D.J.N. Limebeer, Linear Robust Control. Prentice Hall, New Jersey, 1995.

[11] P. Gahinet and P. Apkarian, "A linear matrix inequality approach to H_∞ control," *Int. J. Robust Nonlinear Control*, vol. 4, pp. 421-448, 1994.

[12] T. Iwasaki and R.E. Skelton, "All controllers for the general H_∞ control problem: LMI existence conditions and state space formulas," *Automatica*, vol. 30, pp. 1307-1317, 1994.

5 Robust H-inf Output Feedback Sliding Mode Control With Applications

Robust H_∞ Output Feedback Sliding Mode Control With Applications

Dr.-Ing Bader W. Juma and Prof. Dr. Herbert Werner

Institute of Control Systems, Hamburg University of Technology
Email:baderwady@yahoo.com, h.werner@tu-harburg.de
Telephone: +49-40-42878-3215, Fax: +49(0)40 42 878-2112

***Abstract*—This paper presents an output feedback sliding mode control scheme for uncertain dynamical systems. The design problem is solved in two steps, involving first a state feedback and then an output feedback problem. First, using the null space dynamics, the sliding surface for the unmatched uncertainty is designed. Then, by tuning the sliding surface a robust controller is constructed for the whole uncertainty; this problem takes the form of static output feedback. Based on this, a dynamic output feedback controller for the system augmented with the sliding surface is designed. The synthesis involves the solution of an LMI and a BMI problem; the BMI problem is solved iteratively. The proposed approach is illustrated by applying it to a well-known robust benchmark problem, and also experimentally on a spring mass system with variable stiffness. Simulation and experimental results show that the proposed method outperforms previous approaches in terms of robust performance.**

I. Introduction

Ideal sliding mode control is a technique that has attracted the attention of researchers for more than four decades because of its attractive theoretical robustness properties [8], [3]. A serious limitation in this technique and most of the literature in this field is that only state feedback is considered, i.e. all states are assumed to be available for measurement. In [10], static output feedback sliding mode control is introduced to nominal system without uncertainty, the necessary and sufficient condition to existence problem is given in [7], [14]. In [7] introduced sliding mode control to system with matched uncertainty, constructed the controller in LMI framework in presence of matched uncertainty using only equivalent control is done by [14]. The controller in [14] is extended by [6] using equivalent control and applied control. In static output feedback sliding mode controller, if the number of system output and input are equal, it is not possible to build sliding hyperplane robust to the unmatched uncertainty unless the system inherent robust stable against unmatched uncertainty. In other words, there is enough stable zero dynamic in the system covering the dynamic of the sliding hyperplane and inherently stable against unmatched uncertainty. Hence, [6] is extended by [22] to design robust sliding hyperplane against matched and unmatched uncertainty by considering more system output than input is available. The sliding hyperplane is designed to minimize the unmatched by formulate it in polytopic formula, the equivalent control part divided to two parts, a part minimize the unmatched uncertainty and part minimize the matched one. The design of static output feedback sliding mode control in [6], [22] is encounter non-convex matrix inequality that non-iterative LMI-based algorithm proposed in [17], [16] is used to design the controller.

As discussed previously, dynamic output feedback sliding mode control can satisfied only the sufficient condition existence problem. In [11], the controller is designed to nominal system without uncertainty. For system has matched and unmatched uncertainty a robust hyperplane design is introduced [1]. In [9], the sliding hyperplane is constructed using LMI, the design is restricted to system has matched uncertainty, this work is extended by [23] to unmatched uncertainty. In [29], a different approach to design state feedback sliding mode control is introduced, the approach is to design dynamic sliding hyperplane by using compensator, the compensator is designed when system in sliding mode. This approach is extended by [20] to design output feedback dynamic sliding mode control using H_∞ control μ synthesis theory, and LPV sliding mode in [24].

In most applications, only measured system outputs are accessible. A more practical approach is taken in [21], where output feedback sliding mode control is considered. However, it is restricted to a specific type of uncertainty, the control action tends to have high magnitude and frequency contents. The design procedure in [5] introduces a general way of selecting a sliding surface for a given system in an optimal way, such that the performance of the reduced order system is balanced against the control costs. In this paper, we propose a new approach to design output feedback sliding mode control scheme where the dynamics on the sliding surface are selected to minimize a H_∞ cost function. The approach is illustrated with its application to a well-known benchmark problem for robust control, the ACC benchmark problem, and to an experimental spring mass system with variable spring constants. The extension of this paper to robust H_2 is introduced in [19]. In [18] extended to LPV system.

The layout of the paper is as follows. The general dynamic model is first introduced, and a method to design a sliding surface and to synthesize a sliding mode output feedback controller is proposed. To illustrate the approach, in section(IV) the ACC benchmark problem model is summarized. Section(V) shows simulation results and a tuning procedure for the proposed controller. Finally, experimental results with an implementation of the proposed controller and a standard robust

LTI output feedback controller based H_∞ minimization on an experimental spring mass system are presented in section(VI).

II. Problem Description

Consider an uncertain dynamic system of the form

$$\begin{aligned} \dot{x} &= (A+\Delta A)x+(B+\Delta B)u \\ y &= Cx \end{aligned} \tag{1}$$

where $\Delta A, \Delta B$ represents the uncertainty in the matrices A, B respectively. The matrix A can be decomposed as $(A_{um}+A_m)$

$$\begin{aligned} A &= \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}, \quad A_{um} = \begin{bmatrix} A_{um_{11}} & A_{um_{12}} \\ 0 & 0 \end{bmatrix} \\ A_m &= \begin{bmatrix} 0 & 0 \\ A_{m_{21}} & A_{m_{22}} \end{bmatrix}, \quad B = \begin{bmatrix} 0 \\ B_2 \end{bmatrix} \end{aligned} \tag{2}$$

where $x \in R^n$ is the state vector, $u \in R^m$ is the control input, $y \in R^p$ is the output. A_m represents matched uncertainty i.e. $B^T A_m \neq 0$, A_{um} is unmatched uncertainty that mean $A_{um}^T B^\perp \neq 0$, and $A_{11} \in R^{(n-m)\times(n-m)}$, $B_2 \in R^{m\times m}$.

The first objective is to design a sliding surface of the form $\mathscr{S} = \{x : Sx = 0\}$, where $S \in R^{m\times n}$ is a full rank matrix which needs to be designed, so that the associated reduced order sliding mode, when the system states are confined to $\mathscr{S}$, has appropriate dynamics. This is a state feedback problem (all states are available for measurement).

The second objective is to design a sliding surface of the form $\mathscr{S} = \{x : S(C_K x_k + D_K y) = 0\}$, where $x_k \in R^n$, with controller state equation $\dot{x}_k = A_K x_k + B_K y$, $A_K \in R^n, B_K \in R^{n\times p}, C_K \in R^{n\times n}, D_K \in R^{n\times p}$. This problem is output feedback.

III. Controller Design

The generalized plant $P(s)$ for (1) has the state space representation

$$\begin{aligned} \dot{x} &= Ax + B_w w + Bu \\ y &= Cx \\ z &= C_w x + D_u u \\ w &= \Delta z \end{aligned} \tag{3}$$

Perturbation of the nominal plant A, B is expressed via fictitious inputs through B_w and fictitious outputs through C_w, D_u. Where the matrix Δ represents the range of admissible perturbations and is assumed to satisfy $\|\Delta\|_\infty < 1$ at all times, leading to the LFT form

$$\dot{x} = (A + B_w \Delta C_w)x + (B + B_w \Delta D_u)u \tag{4}$$

For the analysis, it is convenient to re-partition the system states x given in (1) to x_1, x_2 where $x_1 \in R^{(n-m)\times 1}$, $x_2 \in R^{m\times 1}$. On the other hand the sliding matrix can be partitioned as $S = \begin{bmatrix} S_1 & S_2 \end{bmatrix}$ where $S_1 \in R^{m\times(n-m)}$, $S_2 \in R^{m\times m}$. To design sliding mode controller, the design procedure will be in two steps. First, find the a part of the sliding surface matrix $S_2^{-1}S_1$ that minimizes the unmatched uncertainty A_{um}, then choose S_2 to minimize the whole uncertainty $\Delta A, \Delta B$. Next, augmented the system with the sliding surface, then design the dynamic part that feeds the controller with the necessary information about the states. The control block diagram is shown in Figure 1.

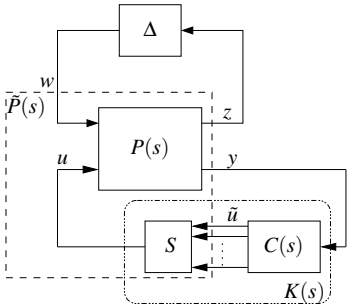


Fig. 1. Closed-loop system with generalized plant

A. State Feedback Sliding Mode Controller

For first order sliding mode control, it is known [5] that when the system is on the sliding surface, the dynamics of the motion are governed by (5),

$$\begin{aligned} \dot{x_1} &= (A_{11} + A_{um_{11}})x_1 + (A_{12} + A_{um_{12}})x_2 \\ x_2 &= -Mx_1 \end{aligned} \tag{5}$$

where $M \in R^{m\times(n-m)}$, and the sliding surface S can be decomposed such that $S = \alpha \begin{bmatrix} M & I_m \end{bmatrix}$. By neglecting $A_{um_{12}}$ the dynamic part of (5) can be written in LFT form as follows

$$\dot{x_1} = (A_{11} + B_{11}\Delta_{11}C_{11})x_1 + A_{12}x_2 \tag{6}$$

The singular value decomposition method of [27] can be used to determine B_{11} and C_{11} such that $\|\Delta_{11}\|_\infty < 1$ for all operating points of the reduced order system. When $A_{um_{12}}$ is not neglected, the system in (5) can be decomposed into a more general LFT from as [4]

$$\dot{x_1} = (A_{11} + B_{11}\Delta_{11}C_{11})x_1 + (A_{12} + B_{11}\Delta_{11}D_{11})x_2 \tag{7}$$

The generalized plant for the reduced order system is shown in Figure 2. The plant variation is represented by an uncertain real gain matrix block Δ_{11} connected between fictitious outputs z_{11} and inputs w_{11}, leading to the state space representation

$$\begin{aligned} \dot{x_1} &= A_{11}x_1 + A_{12}x_2 + B_{11}w_{11} \\ y &= x_1 \\ z_{11} &= C_{11}x_1 + D_{11}x_2 \\ w_{11} &= \Delta_{11}z_{11} \end{aligned} \tag{8}$$

For the state feedback problem, assume the system given in (3) to have all states available for measurement, i.e. $C = I$. The problem considered in this section is to design a sliding surface S, such that the closed loop transfer function T_{zw} satisfies $\|T_{zw}\|_\infty < \gamma$ for all $\|\Delta\|_\infty < 1$. This can be achieved by first designing M to guarantee $\|T_{z_{11}w_{11}}\|_\infty < \gamma_{11}$ for the reduced order system (5). The problem is now to minimize $\|T_{z_{11}w_{11}}\|_\infty$ for the generalized plant in (8). Therefore, M is designed to make $\|T_{z_{11}w_{11}}\|_\infty$ less than γ_{11}. This is a standard LMI problem, i.e. M can be computed using the robust control toolbox for Matlab [2]. On the other hand, finding the value of the scaling parameter α involves solving a BMI problem.

The design problem can now be solved as a static output

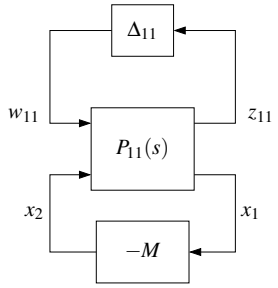


Fig. 2. Configuration with reduced-order plant

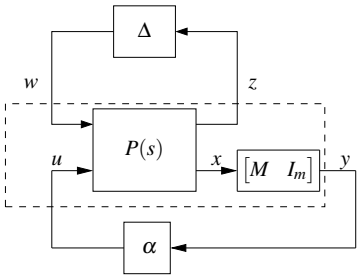


Fig. 3. Static output feedback problem

feedback problem, as shown in Figure 3. Solutions to this problem are discussed e.g. in [13], [15]. For a single-input system, α can be computed from the BMI problem using linear search - fixing it makes it into an LMI problem that can be solved for the value of γ. The best value of α is the one that gives the smallest value for γ.
If there is more than one input to the system, α will be a matrix and a genetic algorithm can be used for the search.

B. Output Feedback Sliding Mode Controller

The problem considered in the second step of the design is to find a dynamic controller $C(s)$ of order n with input y and output $\tilde{u}$

$$C(s) = \left[\begin{array}{c|c} A_K & B_K \\ \hline C_K & D_K \end{array}\right] \tag{9}$$

such that $\|T_{zw}\|_\infty < \gamma$ over all admissible perturbations Δ. The closed-loop block diagram is shown in Figure 4. where the

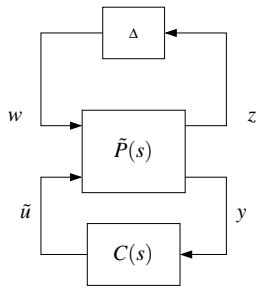


Fig. 4. Dynamic output feedback problem

generalized plant $P(s)$ in Figure 1 is changed to $\tilde{P}(s)$ in Figure 4, and the generalized plant $\tilde{P}(s)$ is given as

$$\begin{aligned} \dot{x} &= Ax + B_w w + B\alpha \begin{bmatrix} M & I_m \end{bmatrix} \tilde{u} \\ y &= Cx \\ z &= C_w x + D_u \alpha \begin{bmatrix} M & I_m \end{bmatrix} \tilde{u} \\ w &= \Delta z \end{aligned} \tag{10}$$

For simplicity suppose $\tilde{D} = D_u \alpha \begin{bmatrix} M & I_m \end{bmatrix}$ and $\tilde{B} = B\alpha \begin{bmatrix} M & I_m \end{bmatrix}$. The closed-loop system from w to z is

$$\begin{aligned} \dot{x_c} &= A_c x_c + B_c w \\ z &= C_c x_c \end{aligned} \tag{11}$$

where

$$\begin{aligned} A_c &= \begin{bmatrix} A + \tilde{B} D_k C & \tilde{B} C_K \\ B_K C & A_K \end{bmatrix}, \qquad B_c = \begin{bmatrix} B_w \\ 0 \end{bmatrix}, \\ C_c &= \begin{bmatrix} C_w + \tilde{D} D_K C & \tilde{D} C_K \end{bmatrix} \end{aligned} \tag{12}$$

The design objective is now to find the controller $C(s)$ that minimizes $\|T_{zw}\|_\infty < \gamma$. The problem has now been converted to an output feedback problem with multi-inputs. This problem can be solved using the robust control toolbox for Matlab.
The controller that will be connected to the system is then

$$K(s) = \left[\begin{array}{c|c} A_K & B_K \\ \hline -\alpha \begin{bmatrix} M & I_m \end{bmatrix} C_K & -\alpha \begin{bmatrix} M & I_m \end{bmatrix} D_K \end{array}\right] \tag{13}$$

The control action that will be applied to the plant is

$$u = k_o \quad sat\left(\frac{S(C_K x_k + D_K y)}{\mu}\right) \tag{14}$$

where $k_o > 0$, and $\mu > 0$ is the width of the boundary layer.

IV. Application to a Benchmark Problem

In this section, the approach presented in the previous section will be illustrated by applying it to a well-known robust benchmark problem, known as ACC benchmark problem, first proposed in [28]. In [25], a score for the achieved performance is defined that will be used here.
Comparison with a standard LTI based H_∞ minimization robust design will show that the proposed sliding mode controller is more robust and achieves a higher performance score than this and other controllers proposed previously for the ACC benchmark problem, see e.g. [25], [12].
The plant to be controlled is a spring mass system, shown in Figure 5, with two masses $m1 = m2 = 1$ connected to each other by a spring with stiffness k, which is uncertain and varies in a given range. A state space model of the system is

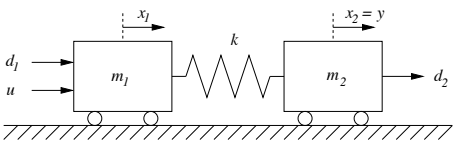


Fig. 5. Two-mass spring system

$$\dot{x} = (A+\Delta A)x + Bu + \begin{bmatrix} G_1 & G_2 \end{bmatrix} \begin{bmatrix} d_1 \\ d_2 \end{bmatrix},$$
$$y = Cx \tag{15}$$

where

$$A+\Delta A = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ -\frac{k}{m_1} & \frac{k}{m_1} & 0 & 0 \\ \frac{k}{m_2} & -\frac{k}{m_2} & 0 & 0 \end{bmatrix}, \quad B = \begin{bmatrix} 0 \\ 0 \\ \frac{1}{m1} \\ 0 \end{bmatrix}$$
$$G = \begin{bmatrix} 0 & 0 \\ 0 & 0 \\ \frac{1}{m1} & 0 \\ 0 & \frac{1}{m2} \end{bmatrix}, \quad C = \begin{bmatrix} 0 & 1 & 0 & 0 \end{bmatrix} \tag{16}$$

Comparing (16) with (15) we obtain

$$A_m = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ -\Delta k & \Delta k & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}, \quad A_{um} = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ \Delta k & -\Delta k & 0 & 0 \end{bmatrix},$$
$$D_m = \begin{bmatrix} 0 \\ 0 \\ \frac{1}{m1} \\ 0 \end{bmatrix} \quad D_{um} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ \frac{1}{m2} \end{bmatrix} \tag{17}$$

where x_1 and x_2 are the position of body 1 and body 2, respectively. x_3 and x_4 are the velocities of body 1 and body 2, respectively. u is the control input acting on body 1. d_1 and d_2 are disturbances acting on body 1 and 2, respectively. y is the measured output. The design requirements is given in [25].

V. Controller Design and Simulation Results

The design procedure divides to two parts. First, design the sliding surface that yields a minimum value of γ when all states are available for feedback. Second, design the dynamic output feedback law and provide the sliding surface with the required states. The sliding surface design requires first to synthesize M that minimize γ_{11} for the reduced order system, while α is used to reduce γ for the full order system. A suitable value for α is chosen by solving the static output feedback problem shown in Figure (3); using linear search. The minimum value of γ for sliding mode state feedback is 1.0692 with α equal 1.35; this will be used in the second design step.

The second design step is to find a dynamic output feedback sliding mode controller $C(s)$, see Figure 4, that minimizes γ. The value of α that has been chosen in the design step for the sliding surface is used when computing $C(s)$. It may however have to be re-tuned, since it had been tuned for a state feedback problem. For the output feedback sliding mode control, the tuning of α can be again considered as static output feedback problem, and solved as a BMI problem. A linear search is used to solve this problem here.

The value of α that minimizes γ for the state feedback problem is not same value as that minimizing γ for the output feedback problem - this value of α is 1.4. The variation of α affects the gain and phase margin: the maximum phase margin is obtained when $\alpha = 0.8$, while the gain margin decreases with increasing α. Increasing the value of α increases the minimum value of the spring constant for which the system is stable, whereas the maximum value of the spring constant for which the system is stable is decreased with increasing α. Thus the range of stability is decreased with increasing α. The effect of α on the robustness measure p_m introduced in [25]: it is obviously decreased with increasing α. Note that the control action in (14) provides a free parameters k_o,μ, that can be tuned to improve the performance. Figures 6-7 show the response of the system to an impulse disturbance applied on mass1 and mass2, respectively, with the tuning parameters α, k_o, μ equal to 0.8, 0.7, and 0.625, respectively. In the reference [25] a scoring

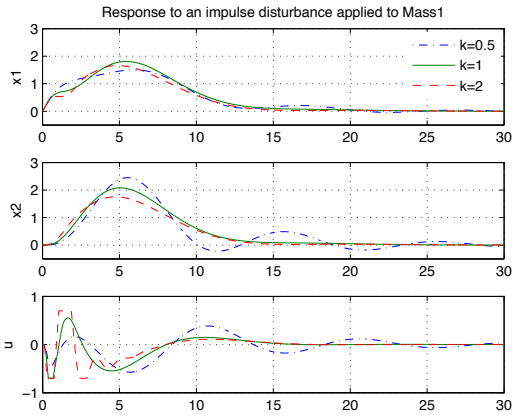


Fig. 6. Impulse disturbance applied at mass1, for α, k_o, μ equal to 0.8, 0.7, 0.625, respectively

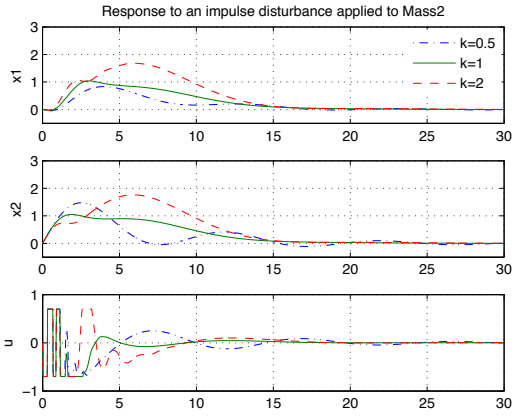


Fig. 7. Impulse applied at mass2, for α, k_o, μ equal to 0.8, 0.7, 0.625, respectively

scheme to evaluate and compare the performance of different controllers for the ACC benchmark problem was proposed. In the same paper, a H_2/classical controller was designed that achieved a score of 7.4. A robust H_2 controller proposed in [12] outperformed this with a score of 8.5, see Table (I). The

sliding mode controller presented above in turn outperforms both controllers and achieves a score of 10.1, see Table (II).

TABLE I
PERFORMANCE MEASURES FOR ROBUST H_2 CONTROLLER [12]

PM	*GM*	t_s(sec)	u_{max}	$k_{min} - k_{max}$	p_m	*Score*
32	6.6	14.5	0.55	0.41-3.1	0.48	8.5

TABLE II
PERFORMANCE MEASURES FOR SLIDING MODE CONTROLLER

PM	*GM*	t_s(sec)	u_{max}	$k_{min} - k_{max}$	p_m	*Score*
37.97	8.5	14.4	0.7	0.24-5.05	0.61	10.1

VI. REAL-TIME CONTROL OF A RECTILINEAR PLANT

In this section, the output feedback sliding mode control scheme is applied to an experimental version of the ACC benchmark problem.

The mechanical plant used for this purpose is the rectilinear mechanism shown in Figure 8. To construct an output feedback

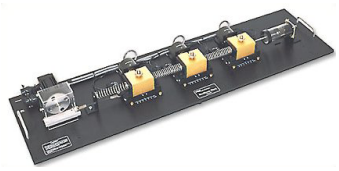


Fig. 8. Experimental version of the ACC benchmark problem

sliding mode controller, the design procedure proposed in section (III) will be used. To tune the controller for performance, an LMI condition representing regional pole constraints is added; here a vertical strip and a conic sector are considered. The fist step in the design is to find M which determines the sliding surface. The LMI condition for M is now combined with the LMI region constraint. In order to convert the design from state feedback to output feedback, the model of the plant in Figure 8 is augmented with the sliding surface and with integral action for accurate tracking.

The augmented system is then used to design the output feedback controller as explained in section (III-B), by adding the LMI region constraint (intersection of vertical strip with conic sector). The tuning parameters are α_l, α_r, θ, the selection of which will be discussed later.

After design the sliding surface and the dynamic part of the sliding mode controller, a first order lag with time constant τ was used as a prefilter to reduce the overshoot. Experimental results show that when the parameters $\alpha_l, \alpha_r, \theta$ are fixed, and the parameter $\alpha_{l_r}, \alpha_{r_r}$ are increased, the steady state error is reduced. However, in general increasing these parameters amplifies noise in the system, a reasonable trade-off used in all results shown in this section is $\alpha_{l_r} = 7$, $\alpha_{r_r} = 3$.

Oscillation in the response can decreased by increasing the conic sector angle θ. The values of the parameter $\alpha_l, \alpha_r, \theta$ for the experimental results shown here are 8.3, 4.3 and 35, respectively.

First, the control law (14) is used for the nominal system (medium stiffness) with k_o, μ equal to 1. Unfortunately, the value of k_o can not be increased beyond one, and the boundary layer μ cannot be decreased below one for the above control law, which means that simply a linear controller is implemented. To reduce the boundary layer μ, the control scheme suggested in [26] turned out to work well in this application. The control law then becomes

$$u = Mk_o(e) \quad sat(\frac{S(C_K x_k + D_K y)}{\mu}) \tag{18}$$

where $Mk_o(e) = k_o \quad M_o \quad |e|$, e is the difference between desired input and controlled output. Using control law (18), the boundary layer can be reduced to $\mu = .0067$ and $k_o = .2$. Experimental results show that the overshoot is increased by increasing k_o, and the reverse is true for the steady state error when the control law (14) is used. On other hand, using the modified control law in (18), the overshoot is decreased and rise time is increased by increasing M_o for the medium and spring. It is clear that the performance is improved when the modified control action in (18) is used. The increasing of M_o reduces the overshoot and increases the rise time for medium and hard stiffness, while for weak stiffness, the overshoot is increased with increasing M_o. By modifying the control law (18), the steady state error can be reduced further: the required modification is as follows:

$$\begin{aligned} if \quad |e| < \varepsilon, \qquad & Mk_o(e) = k_o \quad M_1 \quad |e| \\ otherwise, \qquad & Mk_o(e) = k_o \quad M_2 \quad |e| \end{aligned} \tag{19}$$

The modified control law (19) involves two gains (M_1, M_2) instead of one gain M_o in control law (18). The experimental results show that increasing M_1 reduces the steady state error, increases oscillation, and by reducing ε the oscillation can be reduced. Figure (9) shows the response for different springs using the proposed control law (19), which improves the steady state error. Finally, for comparison a standard robust LTI output

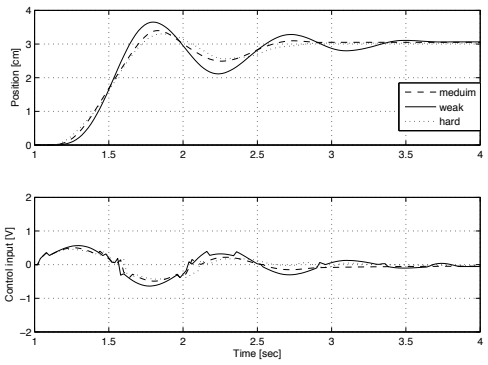


Fig. 9. Experimental results: response with $M_1 = 4, M_2 = 2.5$, $\varepsilon = 0.5$ for three different springs

feedback controller based H_∞ minimization is designed using the same pole constraints that were used for the sliding mode

controller. The system response with this controller is shown in Figure 10. It is obvious from the comparison of the response of the two controllers in Figures 9 - 10, the sliding mode controller outperforms the standard robust LTI controller based H_∞.

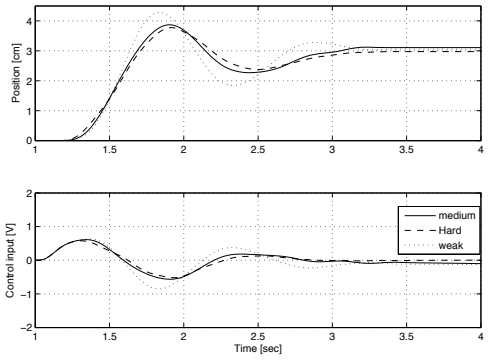


Fig. 10. Experimental results: response for robust LTI output feedback control based H_∞ for three springs

VII. Conclusion

A systematic design and tuning strategy for H_∞ output feedback sliding mode controllers has been proposed. The first step is to design a robust sliding surface that minimizes the H_∞ norm for a part of the unmatched uncertainty or the all unmatched uncertainty. Secondly, the system is augmented with the sliding surface, then one designs an output feedback dynamic sliding mode controller that is robust against the matched and unmatched uncertainty. The proposed control scheme has been tested on a benchmark problem and compared with previously published results as well as a standard robust LTI based H_∞ minimization controller. Simulation results show that the proposed controller achieves a score is higher than that of previous controllers: The sliding mode control score is 10.1, while for the Classical/H_2 and robust H_2 scheme it is 7.4 and 8.5, respectively. In addition, the proposed controller has been applied successfully to an experimental version of the benchmark problem. Again, comparison with a robust LTI controller based H_∞ shows that sliding mode control gives less overshoot, better steady state error and faster rising time.

References

[1] S. K. Bag, S. K. Spurgeon, and C. Edwards. Output feedback sliding mode design for linear uncertain systems. *Control Theory and Applications, IEE Proceedings* -, 144:209–216, May 1997.

[2] G. Balas, R. Chiang, A. Packard, and M. Safonov. *Robust control toolbox 3*. The MathWorks Inc., 2009.

[3] B. Drazenvic. The invariance condtions in variable structure systems. *Automatica*, 5:287–295, 1969.

[4] M. Durrant. *VSC-HVDC power transmission control: a case study in linear controller design for uncertain nonlinear plants*. PhD thesis, Institute of Control System, Hamburg University of Technology, 2007.

[5] C. Edwards. A practical method for the design of sliding mode controllers using linear matrix inequalities. *Automatica*, 40:1761–1769, 2004.

[6] C. Edwards, A. Akoachere, and S. K. Spurgeon. Sliding-mode output feedback controller design using linear matrix inequalities. *IEEE Transactions on Automatica Control*, 46(1):115–119, January 2001.

[7] C. Edwards and S. K. Spurgeon. Sliding mode stabilization of uncertain systems using only output information. *International Journal of Control*, 62(5):1129–1144, 1995.

[8] C. Edwards and S. K. Spurgeon. *Sliding Mode Control: Theory and Applications*. Taylor and Francis, 1998.

[9] C. Edwards and S. K. Spurgeon. Linear matrix inequality methods for designing sliding mode output feedback controllers. *Control Theory and Applications, IEE Proceedings* -, 150(5), Sep. 2003.

[10] R. El-Khazali and R. A. Decarlo. Variable structure output feedback control. pages 871–875. American Control Conference, Chicago, III, 1992.

[11] R. El-Khazali and R. A. Decarlo. Output feedback variable structure control design using dynamic compensation for linear systems. pages 954–958. American Control Conference, San Francisco, CA, USA, 1993.

[12] A. O. Farag. *A practical approach to robust control using linear matrix inequalities*. PhD thesis, Institute of Control System, Hamburg University of Technology, 2005.

[13] J. C. Geromel, Bernussou, and P. L. D. Peres. Decentralized control through parameter space optimization. *Automatica*, 30(10):1565–1578, 1994.

[14] G. Gu. Stabilizability conditions of multivariable uncertain systems via output feedback control. *IEEE Transactions on Automatic Control*, 35(8):924–927, 1990.

[15] T. Iwasaki and R. E. Skelton. Parameterization of all stabilizing controllers via quadratic Lyapunov functions. *Journal of Optimization Theory and Applications*, 85(2):291–307.

[16] R. E. Benton Jr. and D. Smith. Static-output feedback stabilization with prescribed degree of stability. *IEEE Transactions on Automatica Control*, 43:1493–1496, 1998.

[17] R. E. Benton Jr. and D. Smith. A non-iterative LMI based algorithm for robust static output feedback stabilization. *International Journal of Control*, 72(14):1322–1330, 1999.

[18] B. W. Juma. *A practical approach to robust sliding mode control*. PhD thesis, Institute of Control System, Hamburg University of Technology, 2010.

[19] B. W. Juma and H. Werner. Mixed H_2/H_∞ robust output feedback sliding mode control. *Submitted to Journal of Optimal Control, Applications and Methods*, 2010.

[20] K. Nonami, H. Nishimura, and H. Tian. H_∞/μ control-based frequency-shaped sliding mode control for flexible structures. *JSME Int. J.,ser. C*, 39(3):493–501, 96.

[21] K.-K. Shyu, Y.-W. Tsai, Y. Yu, and K.-C. Chang. Dynamic output feedback sliding mode design for a class of linear unmathced uncertainty systems. *International Journal of Control*, 73(16):1463–1474, 2000.

[22] J.M. Andrade-Da Silva, S. K. Spurgeon, and C. Edwards. Sliding mode output feedback controller design using linear matrix inequalities with application to aircraft systems. pages 227–232. International Workshop on Variable Structure Systems, Antalya, June 2008.

[23] J.M.A.-D Silva, C. Edwards, and S. K. Spurgeon. Linear matrix inequality based dynamic output feedback sliding mode control for uncertain plants. In *American Control Conference, St. Louis, MO*, pages 763–768, June 2009.

[24] S. Sivrioglu and K. Nonami. Sliding mode control with time-varying hyperplane for AMB systems. *IEEE/ASME Transactions on Mechatronics*, 3(1), March 1998.

[25] P. M. Thompson. Classical/H_2 solution for a robust control design benchmark problem. *Journal of Guidance, Control and Dynamics*, 18(11):546–555, 1995.

[26] V. Utkin and H. Lee. Chattering problem in sliding mode control systems. *International Workshop on Variable Structure Systems,Alghero, Italy*, June 5-7 2006.

[27] H. Werner, P. Kobra, and Tai chen Yang. Robust tunning of power system stabilizers using *LMI*-techniques. *IEEE Transactions on Control Systems Technology*, 11:147–52, 2003.

[28] B. Wie and D. S. Bernstein. A benchmark problem for robust control design. pages 961–962. American Control Conference, San Diego, CA, USA, 1990.

[29] D. Young and U. Ozguner. Frequency shaping compensator design for sliding mode control. *International Journal of Control*, 57(5), 1993.

6 Quadratic estimation of discrete-time signals using observations with multiple packet dropouts

Quadratic estimation of discrete-time signals using observations with multiple packet dropouts

R. Caballero-Águila
Dpto. de Estadística e I.O.
Universidad de Jaén,
Paraje Las Lagunillas s/n,
23071 Jaén, Spain
Email: raguila@ujaen.es

A. Hermoso-Carazo and J. Linares-Pérez
Dpto. de Estadística e I.O.
Universidad de Granada,
Avda. Fuentenueva s/n,
18071 Granada, Spain
Email: ahermoso@ugr.es, jlinares@ugr.es

***Abstract*—The least-squares quadratic filtering and fixed-point smoothing problems of discrete-time stochastic signals from observations with multiple packet dropouts are addressed. The random dropouts are modelled by introducing a sequence of Bernoulli random variables with known distributions in the observation model. A recursive estimation algorithm is deduced without requiring full knowledge of the state-space model generating the signal process, but only information about the dropout probabilities and the moments of the processes involved. Defining suitable augmented signal and observation vectors, the quadratic estimation problem is reduced to the linear estimation problem of the augmented signal based on the augmented observations, which is solved by using an innovation approach.**

I. Introduction

Over the past few years, research on networked systems has gained lot of interest. Classical estimation methods are not appropriate for these systems in which time delay and/or data packet dropouts are unavoidable due to numerous causes, such as network congestion, random failures in the transmission mechanism, accidental loss of some measurements, or data inaccessibility at certain times.

Under the assumption that the state-space model of the signal to be estimated is known, several modifications of conventional linear estimation algorithms have been proposed to incorporate the effects of random delays on the measurement arrival (see e.g. Ray et al. [1]). Also, many results have been reported on linear estimation for systems with packet dropouts (see e.g. Sahebsara et al. [2] and Sun et al. [3] for systems with multiple packet dropouts, and Sun [4] for the case when the number of consecutive packet dropouts is bounded by a finite number). Nevertheless, in some practical situations the state-space model of the signal is not available and another type of information, for example about the covariance functions of the processes involved in the observation equation, must be processed for the estimation. In this context, linear estimation algorithms from randomly delayed observations based on covariance information have been derived, for example, in Nakamori et al. [5] and, also, quadratic estimators, which improve significantly the performance of linear ones, have been proposed in Hermoso and Linares [6], among others. However, for systems with packet dropouts, estimation problems using covariance information have not been well studied yet.

In this paper a least-squares quadratic filtering and fixed-point smoothing algorithm is proposed for observation models with multiple packet dropouts. The measurement packet dropouts are modelled by introducing a sequence of Bernoulli random variables, whose values (one or zero) indicate if the current measure is available or lost (in which case, the latest measurement is processed instead). For the quadratic estimation approach the signal and observation vectors are augmented by assembling the original vectors with their second-order powers. Then, by using an innovation approach, the linear estimator of the augmented signal based on the augmented observations is obtained, providing the required quadratic estimator.

The rest of the paper is organized as follows. In Section II the observation model considered and the hypotheses on the signal and noise processes are presented. The least-squares quadratic estimation problem is formulated in Section III, where the augmented observation model and the statistical properties of the augmented vectors are established; the quadratic estimation problem is then reduced to the linear estimation problem of the augmented signal, and the innovation technique used to address such linear estimation problem is described. The least-squares quadratic estimation algorithm is derived in Section IV which includes recursive formulas to obtain the estimation error covariance matrices; these matrices provide a global measure of the estimators accuracy. Finally, in Section V, a numerical simulation example is presented to show the effectiveness of the estimation algorithm proposed in the current paper, and some conclusions are drawn in Section VI.

II. Observation model

Consider an n-dimensional signal vector, z_k, whose measured output at the sampling time k, denoted by $\widetilde{y}_k$, is perturbed by an additive noise vector v_k; that is,

$$\widetilde{y}_k = z_k + v_k, \qquad k \geq 1. \tag{1}$$

Assume that, at the initial time $k = 1$, the measured output $\widetilde{y}_1$ is always available and, hence, the measurement processed for the estimation is equal to the real measurement, $y_1 = \widetilde{y}_1$. However, at any time $k > 1$, the measured output $\widetilde{y}_k$ can

be randomly dropped-out during network transmission, which occurs with known probability. Consequently, the measurement processed at time k, y_k, will be either the current measured output $\widetilde{y}_k$ (with probability p_k) or, if such measured output is lost during transmission, then the latest measurement received y_{k-1} will be processed (with probability $1-p_k$); that is

$$y_k = \begin{cases} \widetilde{y}_k, & \text{with probability } p_k \\ y_{k-1}, & \text{with probability } 1-p_k. \end{cases}$$

Therefore, the following model for the measurements processed to estimate the signal is considered:

$$y_k = \gamma_k \widetilde{y}_k + (1-\gamma_k) y_{k-1}, \quad k>1; \quad y_1 = \widetilde{y}_1, \qquad (2)$$

where $\{\gamma_k;\ k>1\}$ is a sequence of mutually independent Bernoulli random variables with $P[\gamma_k = 1] = p_k$.

This model, considered by Sahebsara et al. [2], can describe multiple packet dropouts if successive values of the Bernoulli variables are equal to zero. For example, if $\gamma_k = \gamma_{k-1} = \gamma_{k-2} = 0$ and $\gamma_{k-3} = 1$, three consecutive data are lost and the measurement received at $k-3$ will be used at $k-2$, $k-1$ and k.

To address the quadratic estimation problem of the signal the following assumptions are needed:

(H1) The signal process $\{z_k;\ k \geq 1\}$ has zero mean and autocovariance function $K_{k,s} = E[z_k z_s^T] = A_k B_s^T,\ s \leq k$, where A and B are $n \times M$ known matrix functions. The autocovariance function of the second-order powers (defined by the Kronecker product, $z^{[2]} = z \otimes z$), $K_{k,s}^{z^{[2]}} = E\left[(z_k^{[2]} - E[z_k^{[2]}])(z_s^{[2]} - E[z_s^{[2]}])^T\right]$ is also factorized as $K_{k,s}^{z^{[2]}} = a_k b_s^T, \quad s \leq k$, where a and b are $n^2 \times L$ known matrix functions. Moreover,

$$K_{k,s}^{z z^{[2]}} = E\left[z_k z_s^{[2]T}\right] = \begin{cases} \alpha_k \beta_s^T, & s \leq k, \\ \varepsilon_k \delta_s^T, & k \leq s, \end{cases}$$

where α, β, ε and δ are $n \times N$, $n^2 \times N$, $n \times P$ and $n^2 \times P$ known matrix functions, respectively.

(H2) The noise process $\{v_k;\ k \geq 1\}$ is a zero-mean white sequence and its moments, up to the fourth one, are also known and denoted as follows

$$R_k^v = E[v_k v_k^T], \quad R_k^{v v^{[2]}} = E[v_k v_k^{[2]T}],$$
$$R_k^{v^{[2]}} = E\left[(v_k^{[2]} - E[v_k^{[2]}])(v_k^{[2]} - E[v_k^{[2]}])^T\right].$$

(H3) The noise $\{\gamma_k;\ k>1\}$ is a sequence of independent Bernoulli random variables with known probabilities $P[\gamma_k = 1] = p_k$.

(H4) The signal $\{z_k;\ k \geq 1\}$ and the noises $\{v_k;\ k \geq 1\}$ and $\{\gamma_k;\ k>1\}$ are mutually independent processes.

III. Quadratic estimation problem

Given the observation model (1)-(2) under assumptions (H1)-(H4), the problem is to find the least-squares (LS) quadratic estimator, $z_{k/L}^Q$, of the signal, z_k, when information on the measurement history up to the L^{th} instant, $\{y_1, \ldots, y_L\}$, is available. More specifically, our aim is to derive recursive algorithms for the filter, $z_{k/k}^Q$, and the estimators $z_{k/L}^Q$ at the fixed-point k, for any $L > k$, that is, the fixed-point smoother.

From hypotheses (H1) and (H2), $E[y_i^{[2]T} y_i^{[2]}] < \infty$, so the required quadratic estimator $z_{k/L}^Q$ exists and can be obtained as the orthogonal projection of z_k on the space of n-dimensional linear transformations of $y_1, \ldots, y_L$ and $y_1^{[2]}, \ldots, y_L^{[2]}$. The technique used to obtain this estimator consists of augmenting the signal and observation vectors by assembling the original vectors and their second-order powers,

$$\mathcal{Z}_k = \begin{pmatrix} z_k \\ z_k^{[2]} \end{pmatrix}, \quad \widetilde{\mathcal{Y}}_k = \begin{pmatrix} \widetilde{y}_k \\ \widetilde{y}_k^{[2]} \end{pmatrix}, \quad \mathcal{Y}_k = \begin{pmatrix} y_k \\ y_k^{[2]} \end{pmatrix}$$

thus deriving the estimator $z_{k/L}^Q$ as the vector constituted of the first n entries of the LS linear estimator of $\mathcal{Z}_k$ based on $\mathcal{Y}_1, \ldots, \mathcal{Y}_L$, whose existence is guaranteed from (H1).

In the next section the relation between the augmented vectors is studied and their statistical properties are analyzed.

A. Augmented observation model

By using (1) and the Kronecker product properties, the following expression for $\widetilde{y}_k^{[2]}$ is obtained

$$\widetilde{y}_k^{[2]} = z_k^{[2]} + (I_{n^2} + K_{n^2})(z_k \otimes v_k) + v_k^{[2]}, \quad k \geq 1,$$

where I_{n^2} is the $n^2 \times n^2$ identity matrix and K_{n^2} is the $n^2 \times n^2$ commutation matrix, which satisfies $K_{n^2}(z_k \otimes v_k) = v_k \otimes z_k$. Hence, the centered augmented vectors $Z_k = \mathcal{Z}_k - E[\mathcal{Z}_k]$ and $\widetilde{Y}_k = \widetilde{\mathcal{Y}}_k - E[\widetilde{\mathcal{Y}}_k]$ satisfy

$$\widetilde{Y}_k = Z_k + V_k, \quad k \geq 1,$$

where

$$V_k = \begin{pmatrix} v_k \\ (I_{n^2} + K_{n^2})(z_k \otimes v_k) + v_k^{[2]} - vec(R_k^v) \end{pmatrix}.$$

Next, using (2) and taking into account that $\gamma_k^2 = \gamma_k$, the following expression for $y_k^{[2]}$ is obtained

$$y_k^{[2]} = \gamma_k \widetilde{y}_k^{[2]} + (1-\gamma_k) y_{k-1}^{[2]}, \quad k>1; \quad y_1^{[2]} = \widetilde{y}_1^{[2]}.$$

Using again the Kronecker product properties and the model hypotheses, it is deduced that the centered augmented vector $Y_k = \mathcal{Y}_k - E[\mathcal{Y}_k]$ satisfy the following equation:

$$Y_k = \gamma_k \widetilde{Y}_k + (1-\gamma_k) Y_{k-1} + (\gamma_k - p_k) C_k,\ k>1; \quad Y_1 = \widetilde{Y}_1,$$

where $C_k = E[\widetilde{\mathcal{Y}}_k] - E_{k-1},\ k>1$, with $E_k = E[\mathcal{Y}_k]$ being recursively calculated from

$$E_k = (1-p_k) E_{k-1} + p_k E[\widetilde{\mathcal{Y}}_k], \quad k \geq 2; \quad E_1 = E[\widetilde{\mathcal{Y}}_1]$$

and

$$E[\widetilde{\mathcal{Y}}_k] = \begin{pmatrix} 0 \\ vec(A_k B_k^T + R_k^v) \end{pmatrix}, \quad k \geq 1.$$

B. Statistical properties of augmented vectors

Clearly, the signal and noise processes, $\{Z_k;\ k \geq 1\}$ and $\{V_k;\ k \geq 1\}$, have zero mean. Their second-order statistical properties, which are derived from (H1)-(H4), are established in the following propositions.

Proposition 1. If the signal process $\{z_k;\ k \geq 1\}$ satisfies (H1), the autocovariance function of the augmented signal process $\{Z_k;\ k \geq 1\}$ can be expressed in a semi-degenerate kernel form; namely,

$$K^Z_{k,s} = E[Z_k Z_s^T] = \mathcal{A}_k \mathcal{B}_s^T, \quad s \leq k,$$

where

$$\mathcal{A}_k = \begin{pmatrix} A_k & \alpha_k & 0_{n\times P} & 0_{n\times L} \\ 0_{n^2\times M} & 0_{n^2\times N} & \delta_k & a_k \end{pmatrix},$$

$$\mathcal{B}_k = \begin{pmatrix} B_k & 0_{n\times N} & \varepsilon_k & 0_{n\times L} \\ 0_{n^2\times M} & \beta_k & 0_{n^2\times P} & b_k \end{pmatrix}.$$

Proposition 2. Under (H1)-(H4), the noise $\{V_k;\ k \geq 1\}$ is a sequence of mutually uncorrelated random vectors with covariance matrices given by

$$E[V_k V_k^T] = R_k^V = \begin{pmatrix} R_k^v & R_k^{vv^{[2]}} \\ R_k^{vv^{[2]}T} & R_k^{22} \end{pmatrix}$$

where

$$R_k^{22} = (I_{m^2} + K_{m^2}) \left(A_k B_k^T \otimes R_k^v\right) (I_{m^2} + K_{m^2}) + R_k^{v^{[2]}}.$$

Moreover, $\{V_k;\ k \geq 1\}$ is uncorrelated with the process $\{Z_k;\ k \geq 1\}$.

C. Linear estimation of the augmented signal Z_k

As indicated previously, to obtain the LS quadratic estimators of the signal, z_k, based on the observations $\{y_1, \ldots, y_L\}$, we consider the LS linear estimation problem of the augmented signal, Z_k, based on the augmented observations $\{Y_1, \ldots, Y_L\}$. This problem is addressed via an innovation approach, which simplifies considerably the derivation of the filtering algorithm, since the innovations constitute a white process.

Let $\nu_i = Y_i - \widehat{Y}_{i/i-1}$, where $\widehat{Y}_{i/i-1}$ denotes the LS linear estimator of Y_i based on the previous observations, $Y_1, \ldots, Y_{i-1}$. For each i, ν_i may be regarded as a measure of the new information or the *innovation* provided by the observation Y_i. It is known that the innovations $\{\nu_i,\ i \leq L\}$ can be determined from the observations $\{Y_i,\ i \leq L\}$ by means of a causal and causally invertible linear transformation. Therefore, each set can be replaced by the other with no loss of information and, consequently, the LS linear filter of the signal Z_k based on the observations $Y_1, \ldots, Y_L$, which is denoted by $\widehat{Z}_{k/L}$, is equal to the LS linear estimator given the innovations $\nu_1, \ldots, \nu_L$. Since the innovations constitute a white process, from the Orthogonal Projection Lemma it is easily proven that the filter is given by

$$\widehat{Z}_{k/L} = \sum_{i=1}^{L} S_{k,i} \Pi_i^{-1} \nu_i, \tag{3}$$

where $S_{k,i} = E[Z_k \nu_i^T]$ and $\Pi_i = E[\nu_i \nu_i^T]$.

Thus, as the estimators are expressed in terms of the innovations, we start by determining them or, equivalently, the predictors $\widehat{Y}_{i/i-1}$. It is clear that $\widehat{Y}_{1/0} = 0$ and, taking into account the model hypotheses, the Orthogonal Projection Lemma leads to

$$\widehat{Y}_{i/i-1} = p_i \widehat{Z}_{i/i-1} + (1 - p_i) Y_{i-1}, \quad k \geq 2. \tag{4}$$

IV. Quadratic filtering and fixed-point smoothing algorithm

Using the properties of the augmented processes, as established in propositions 1 and 2, we derive recursive algorithms for the linear filtering and fixed-point smoothing estimators, $\widehat{Z}_{k/L},\ L \geq k$, of the augmented signal Z_k. These estimators allow us to obtain the required quadratic filtering and fixed-point smoothing estimators of the original signal z_k, just by extracting the first n entries.

Theorem 1. The quadratic filtering and fixed-point smoothing estimators, $z^Q_{k/L},\ L \geq k$, of the signal z_k are given by

$$z^Q_{k/L} = \Upsilon \widehat{Z}_{k/L}, \quad L \geq k$$

where Υ is the operator which extracts the first n entries of $\widehat{Z}_{k/L}$, the linear estimators of the augmented signal Z_k, which are recursively obtained by

$$\widehat{Z}_{k/L} = \widehat{Z}_{k/L-1} + S_{k,L} \Pi_L^{-1} \nu_L, \qquad L > k \tag{5}$$

from the initial condition

$$\widehat{Z}_{k/k} = \mathcal{A}_k O_k. \tag{6}$$

The innovation, ν_L, satisfies

$$\begin{array}{l} \nu_L = Y_L - p_L \mathcal{A}_L O_{L-1} - (1 - p_L) Y_{L-1}, \quad L \geq 2; \\ \nu_1 = Y_1, \end{array} \tag{7}$$

where the vectors O_L are recursively calculated from

$$O_L = O_{L-1} + J_L \Pi_L^{-1} \nu_L, \quad L \geq 1; \qquad O_0 = 0. \tag{8}$$

The matrix function J satisfies

$$J_L = p_L \left[\mathcal{B}_L^T - r_{L-1} \mathcal{A}_L^T\right], \quad L \geq 2; \quad J_1 = \mathcal{B}_1^T, \tag{9}$$

where r_L are recursively obtained from

$$r_L = r_{L-1} + J_L \Pi_L^{-1} J_L^T, \quad L \geq 1; \qquad r_0 = 0. \tag{10}$$

The covariance matrix of the innovation, Π_L, verifies

$$\begin{array}{l} \Pi_L = \Sigma_L^Y - p_L^2 \mathcal{A}_L r_{L-1} {\mathcal{A}_L}^T - (1 - p_L)^2 \Sigma_{L-1}^Y \\ \qquad -p_L(1 - p_L) \left[\mathcal{A}_L G_{L-1} + G_{L-1}^T {\mathcal{A}_L}^T\right], \quad L \geq 2; \\ \Pi_1 = \Sigma_1^Y, \end{array} \tag{11}$$

where Σ_L^Y and G_L are recursively calculated from

$$\begin{array}{l} \Sigma_L^Y = p_L(\mathcal{A}_L \mathcal{B}_L^T + R_L^V) + (1 - p_L) \Sigma_{L-1}^Y \\ \qquad +p_L(1 - p_L) C_L C_L^T, \quad L \geq 2; \\ \Sigma_1^Y = \mathcal{A}_1 \mathcal{B}_1^T + R_1^V \end{array} \tag{12}$$

and

$$\begin{array}{l} G_L = J_L + p_L r_{L-1} \mathcal{A}_L^T + (1 - p_L) G_{L-1}, \ L \geq 1; \\ G_0 = 0. \end{array} \tag{13}$$

Finally, the matrices $S_{k,L}$ are calculated from

$$\begin{array}{l} S_{k,L} = p_L[\mathcal{B}_k - H_{k,L-1}]\mathcal{A}_L^T, \quad L > k, \\ S_{k,k} = \mathcal{A}_k J_k \end{array} \tag{14}$$

where $H_{k,L}$ satisfy

$$\begin{array}{l} H_{k,L} = H_{k,L-1} + S_{k,L}\Pi_L^{-1} J_L^T, \qquad L > k, \\ H_{k,k} = \mathcal{A}_k r_k. \end{array} \tag{15}$$

Proof. From the general expression (3), it is clear that the linear fixed-point smoothers of the signal Z_k are recursively obtained by relation (5), and its initial condition is obviously provided by the filter, $\widehat{Z}_{k/k}$.

From (3), in order to determine the filter, the coefficients $S_{k,i} = E[Z_k \nu_i^T]$, must be calculated for $i \leq k$. Using expression (4) in $\nu_i = Y_i - \widehat{Y}_{i/i-1}$ yields

$$\begin{array}{l} S_{k,i} = E[Z_k Y_i^T] - p_i E[Z_k \widehat{Z}_{i/i-1}^T] - (1-p_i)E[Z_k Y_{i-1}], \ i \geq 2 \\ S_{k,1} = E[Z_k \widetilde{Y}_1^T]. \end{array}$$

Again, taking into account (3) for the predictors, $\widehat{Z}_{i/i-1}$, and using that $E[Z_k \widetilde{Y}_i^T] = \mathcal{A}_k \mathcal{B}_i^T$ for $1 \leq i \leq k$, since $E[Z_k \nu_j^T] = S_{k,j}$, we have

$$\begin{array}{l} S_{k,i} = p_i \mathcal{A}_k \mathcal{B}_i^T - p_i \displaystyle\sum_{j=1}^{i-1} S_{k,j} \Pi_j^{-1} S_{i,j}^T, \quad 2 \leq i \leq k \\ S_{k,1} = \mathcal{A}_k \mathcal{B}_1^T. \end{array}$$

This expression for $S_{k,i}$ guarantees that

$$S_{k,i} = \mathcal{A}_k J_i, \qquad 1 \leq i \leq k, \tag{16}$$

where J is a function satisfying

$$\begin{array}{l} J_i = p_i \mathcal{B}_i^T - p_i \displaystyle\sum_{j=1}^{i-1} J_j \Pi_j^{-1} S_{i,j}^T, \quad 2 \leq i \leq k \\ J_1 = \mathcal{B}_1^T. \end{array} \tag{17}$$

Hence, if we denote

$$O_L = \sum_{i=1}^{L} J_i \Pi_i^{-1} \nu_i, \qquad O_0 = 0, \tag{18}$$

which, obviously, satisfies (8), expression (6) for the filter is deduced from (3), (16) and (18) for $L = k$.

Similarly, the one-stage predictors of the signal are given by

$$\widehat{Z}_{L/L-1} = \mathcal{A}_L O_{L-1}, \tag{19}$$

and expression (7) for $\nu_L = Y_L - \widehat{Y}_{L/L-1}$ is obtained by substituting (19) in (4) for $i = L$.

Next, taking into account that, from (16), $S_{L,i} = \mathcal{A}_L J_i$ for $1 \leq i \leq L$, and by denoting

$$r_L = E\left[O_L O_L^T\right] = \sum_{i=1}^{L} J_i \Pi_i^{-1} J_i^T, \quad r_0 = 0, \tag{20}$$

from (17) for $i = L$, expression (9) for J_L is easily derived. The recursive formula (10) for r_L is obvious from (20).

By denoting $\Sigma_L^Y = E[Y_L Y_L^T]$ and $G_L = E[O_L Y_L^T]$, it is clear that the innovation covariance matrix $\Pi_L = E[Y_L Y_L^T] - E[\widehat{Y}_{L/L-1} \widehat{Y}_{L/L-1}^T]$ satisfies expression (11).

Recursive relations (12) and (13) for Σ_L^Y and G_L, respectively, are derived as follows:

- Clearly, from the model hypotheses, $E[\widetilde{Y}_L \widetilde{Y}_L^T] = \mathcal{A}_L \mathcal{B}_L^T + R_L^V$; hence, taking into account that $E[\gamma_L^2] = p_L$, $E[(1-\gamma_L)^2] = 1 - p_L$ and $E[\gamma_L(1-\gamma_L)] = 0$, the expression (12) for Σ_L^Y is obtained.
- We write $G_L = E[O_L Y_L^T] = E[O_L \nu_L^T] + E[O_L \widehat{Y}_{L/L-1}^T]$. Clearly, from (18) and since the innovation is a white process, we have $E[O_L \nu_L^T] = J_L$ and $E[O_L \widehat{Y}_{L/L-1}^T] = E[O_{L-1} \widehat{Y}_{L/L-1}^T]$. Next, from (4), (19) and (20), we have $E[O_{L-1} \widehat{Y}_{L/L-1}^T] = p_L r_{L-1} \mathcal{A}_L^T + (1-p_L) G_{L-1}$, and expression (13) is obtained.

Finally, we must prove (14) for $S_{k,L} = E[Z_k \nu_L^T]$ and (15) for $H_{k,L}$. Using (7) for ν_L, and since $E[Z_k Y_L^T] = p_L \mathcal{B}_k \mathcal{A}_L^T$ for $L > k$, we obtain

$$S_{k,L} = p_L \mathcal{B}_k \mathcal{A}_L^T - p_L E[Z_k O_{L-1}^T] \mathcal{A}_L^T,$$

which leads to (14), just denoting $H_{k,L} = E[Z_k O_L^T]$. The initial condition in (14) is immediately clear from (16).

The recursive relation (15) is derived from (8) and its initial condition is obtained from (6) and (20), taking into account that, from the Orthogonal projection Lemma, $E[Z_k O_k^T] = E[\widehat{Z}_{k/k} O_k^T]$.

A. Error covariance matrices

The performance of the LS estimators $\widehat{Z}_{k/L}$, $L \geq k$, is measured by the covariance matrices of the estimation errors,

$$P_{k/L} = E\left[Z_k Z_k^T\right] - E\left[\widehat{Z}_{k/L} \widehat{Z}_{k/L}^T\right], \quad L \geq k.$$

Using the recursive relation (5) for $\widehat{Z}_{k/L}$, these matrices can be written as

$$P_{k/L} = P_{k/L-1} - S_{k,L} \Pi_L^{-1} S_{k,L}^T, \quad L > k \tag{21}$$

The initial condition is $P_{k/k}$, the error covariance matrix of the filter $\widehat{Z}_{k/k} = \mathcal{A}_k O_k$ which, taking into account that $E[Z_k Z_k^T] = \mathcal{A}_k \mathcal{B}_k^T$ and $r_k = E[O_k O_k^T]$, is given by:

$$P_{k/k} = \mathcal{A}_k \left[\mathcal{B}_k^T - r_k \mathcal{A}_k^T\right].$$

The first $n \times n$ blocks of the matrices $P_{k/L}$, $L \geq k$, constitute the covariance matrices of the quadratic smoothing and filtering errors, thus providing a measure of the accuracy of the respective quadratic estimators.

V. Computer example

In this section a numerical simulation example is shown to illustrate the feasibility and effectiveness of the proposed quadratic estimation algorithm. For this purpose, we have simulated 100 values of the signal to be estimated and the corresponding observations with multiple packet dropouts. Using these observations, both linear and quadratic estimates of the

signal are calculated and the corresponding error covariance matrices are provided to measure the estimation accuracy.

We consider a zero-mean scalar signal $\{z_k;\ k \geq 1\}$ such that the autocovariance and cross-covariance functions of this signal and their second-order powers are given by

$$K^z_{k,s} = 1.025641 \times 0.95^{k-s}, \quad s \leq k,$$
$$K^{z^2}_{k,s} = 2.1038795 \times 0.95^{2(k-s)}, \quad s \leq k,$$
$$K^{zz^2}_{k,s} = 0, \quad \forall s, k;$$

hence, according to hypothesis (H1), the functions which constitute these covariance functions can be defined as follows:

$$A_k = 1.025641 \times 0.95^k, \quad B_k = 0.95^{-k},$$
$$a_k = 2.1038795 \times 0.95^{2k}, \quad b_k = 0.95^{-2k},$$
$$\alpha_k = \beta_k = \varepsilon_k = \delta_k = 0.$$

For the simulations, the signal is assumed to be generated by the following first-order autoregressive model

$$z_{k+1} = 0.95 z_k + w_k$$

where $\{w_k;\ k \geq 1\}$ is a zero-mean white Gaussian noise with $Var\,[w_k] = 0.1$, for all k.

The real measurements of the signal, $\widetilde{y}_k = z_k + v_k$, are perturbed by a white noise, $\{v_k;\ k \geq 1\}$, with distribution

$$P[v_k = -8] = \frac{1}{8}, \quad P\left[v_k = \frac{8}{7}\right] = \frac{7}{8}, \quad \forall k \geq 0;$$

hence,

$$E[v_k] = 0, \quad R^v_k = 9.142857,$$
$$R^{vv^2}_k = -62.693878, \quad R^{v^2}_k = 429.900875.$$

Now, according to our theoretic study, we suppose that, at any sampling time $k > 1$, the measurement processed for the estimation, y_k, can be either the current measured output $\widetilde{y}_k$, with constant probability p, or the latest measurement received y_{k-1}, with probability $1 - p$; that is, the measurements of the signal are given by

$$y_k = \gamma_k \widetilde{y}_k + (1 - \gamma_k) y_{k-1}, \quad k > 1; \quad y_1 = \widetilde{y}_1$$

with

$$\widetilde{y}_k = z_k + v_k, \quad k \geq 1.$$

First, considering a fixed value of the probability $p = 0.5$, the error variances of the linear and quadratic estimators are calculated, allowing us to compare the performance of both estimators. The error variances of the linear and quadratic filters and fixed-point smoothers, are displayed in Figure 1 which shows, on the one hand, that the quadratic estimation error variances are less than the linear ones (confirming the superiority of the quadratic estimators over the linear ones) and, on the other, that the estimation accuracy of the smoothers is superior to that of the filters and, also, that the performance of the fixed-point smoothers improves as the number of available observations increases.

Next, we compare the performance of the estimators considering different values of the probability p. Since the error variances show insignificant variation from the 20th iteration onwards, only the error variances at a specific iteration are considered. In Figure 2 the linear and quadratic filtering and smoothing error variances at $k = 100$ are displayed versus p. This figure shows, as expected, that both, linear and quadratic estimators, have better performance (the error variances are smaller) as the probability p increases or, equivalently, as the dropout probability decreases.

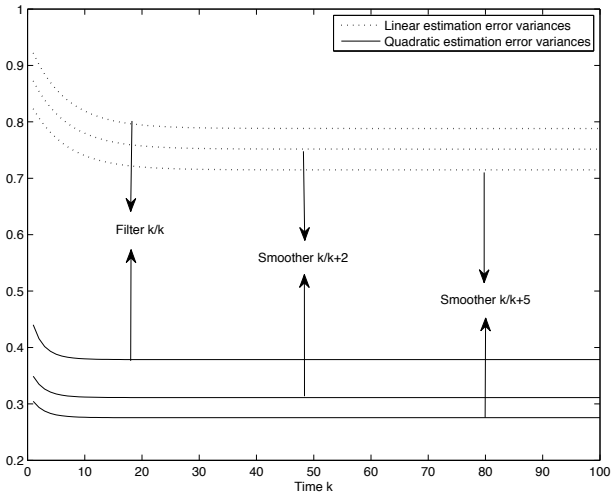


Fig. 1. Linear and quadratic estimation error variances for $p = 0.5$.

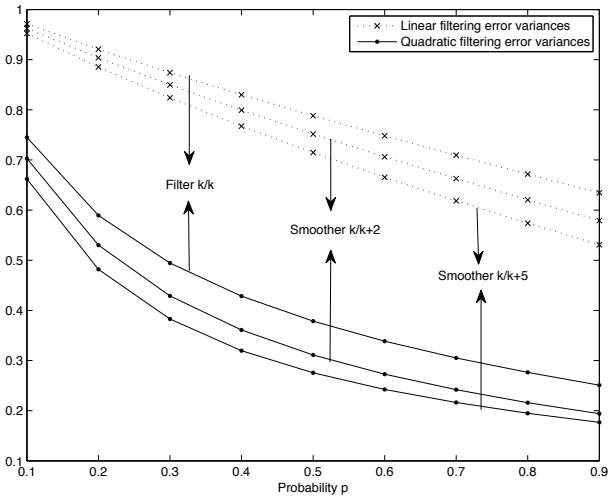


Fig. 2. Linear and quadratic estimation error variances versus p.

Finally, Figure 3 displays 100 simulated values of the signal, together with the linear filtering estimates and the quadratic filtering and fixed-point smoothing estimates, for the value $p = 0.5$. This figure shows, on the one hand, that the quadratic filtering estimates follow the signal evolution better than the linear ones and, on the other, that the signal evolution is followed more accurately by the smoothing estimates, agreeing with the comments made about Figure 1.

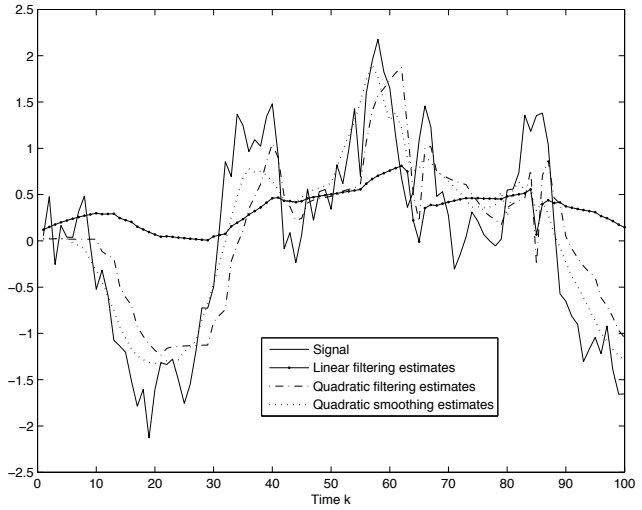


Fig. 3. Simulated signal, linear and quadratic estimates when $p = 0.5$.

VI. Conclusion

A recursive algorithm is proposed for the LS quadratic filter and fixed-point smoother from observations featuring multiple packet dropouts, a realistic assumption in networked stochastic systems where, generally, transmission losses are unavoidable due to the unreliable network characteristics. We assume that, when the current measurement is not available, the latest measurement is processed for the estimation.

To address the quadratic estimation problem, an augmented signal and observation vectors are introduced by assembling the original vectors with their second-order powers, defined by the Kronecker product. Using an innovation approach, the linear estimator of the augmented signal based on the augmented observations is obtained, providing the required quadratic estimator.

The estimation algorithm does not require the knowledge of the signal state-space model, but only the autocovariance and crosscovariance functions of the signal and its second-order powers, and the same information on the additive noise. Furthermore, our estimators only depend on the data arrival probabilities, but do not need to know if a measurement is received or lost at a particular sampling time. To measure the performance of the estimators, recursive formulas for the estimation error covariance matrices are also proposed.

To illustrate the theoretical results established in this paper, a simulation example is presented, showing the feasibility of the proposed algorithm and the superiority of the quadratic estimators over the linear ones.

Acknowledgment

This research is partially supported by Ministerio de Educación y Ciencia (grant No. MTM2008-05567) and Junta de Andalucía (grant No. P07-FQM-02701).

References

[1] A. Ray, L. W. Liou and J. H. Shen, "State estimation using randomly delayed measurements", *Journal of Dynamic Systems, Measurement and Control*, vol. 115, pp. 19–26, 1993.

[2] M. Sahebsara, T. Chen and S. L. Shah, "Optimal H_2 filtering with random sensor delay, multiple packet dropout and uncertain observations", *International Journal of Control*, vol. 80, pp. 292–301, 2007.

[3] S. L. Sun, L. H. Xie, W. D. Xiao and Y. C. Soh, "Optimal linear estimation for systems with multiple packet dropouts", *Automatica*, vol. 44(5), pp. 1333–1342, 2008.

[4] S. Sun, "Linear minimum variance estimators for systems with bounded random measurement delays and packet dropouts", *Signal Processing*, vol. 89, pp. 1457–1466, 2009.

[5] S. Nakamori, R. Caballero-Águila, A. Hermoso-Carazo, and J. Linares-Pérez, "Recursive estimators of signals from measurements with stochastic delays using covariance information", *Applied Mathematics and Computation*, vol. 162, pp. 65–79, 2005.

[6] A. Hermoso-Carazo and J. Linares-Pérez, "Linear and quadratic least-squares estimation using measurements with correlated one-step random delay" *Digital Signal Processing: A Review Journal*, vol. 18 (3), pp. 450–464, 2008.

7 Modelling of A Helicopter System

Modelling of A Helicopter System

K.K.T. Thanapalan
Faculty of Engineering Sciences
University College London
London WC1E 6BT, U.K
E-mail: kary.thanapalan@ucl.ac.uk

***Abstract* — This paper considers modelling and simulation study of a helicopter system – UH-60 Black Hawk helicopter. Mathematical model of single main rotor helicopters is presented in this paper. For the convenience of presentation, force and moment expressions of the various helicopter components are given in the paper to bridge a generic model to the model of UH-60 Black Hawk helicopters. For simulation study a UH-60 like Flightlab GRM model (Generic Rotorcraft Model) is used. Comparisons are made between the simulation results and flight test data. A general agreement exits but where disagreements and anomalies occur, clues are gathered to give explanation. Overall the model represents the UH-60 Black Hawk helicopter. This model can be used for controller development to improve flight handling quality and performances.**

Keywords - modelling; simulation; model validation; helicopter system; flight control

I. INTRODUCTION

A helicopter has six degrees of freedom in its motions: up/down, fore/aft (longitudinal motion), left/right (lateral motion), pitching, rolling, and yawing. The motions of a helicopter are achieved by; 1) collectively changing the pitch of all the main rotor blades, thus increasing rotor thrust (collective pitch); 2) cyclically changing the pitch as a sinusoidal function of azimuth which tilts the tip-path-plane fore/aft or left/right and changes the thrust vector direction (cyclic pitch); and 3) collectively changing the tail rotor pitch, which changes tail rotor thrust and thus the yaw moment. A helicopter pilot must simultaneously control three forces and moments, hence, control of a helicopter, is a difficult task indeed. A helicopter pilot typically has at his disposal a cyclic stick to control both fore/aft motions (pitch control) and left/right motion (roll control), a collective lever to control up and down motions (vertical control), and pedals to control left and right yawing motions (yaw control). Lift, thrust, pitching, and rolling control comes from the main rotor while yawing control comes from the tail rotor (Bramwell, 1976, Stepniewski, *et al,* 1984). Analyse the dynamic problems of controlling a helicopter and to develop control schemes for alleviating these problems it is necessary to derive a dynamic model for helicopters. The dynamic model should be well suited to stability and control analysis, which may involve linearized equations of motion about possible equilibrium positions.

In the following section a mathematical model of a single main rotor helicopter is presented. The forces and moments from the different elements of helicopter are discussed in details. Then the model of UH-60 helicopter has been derived and simulation study has been conducted. The results are presented in this paper.

II. DYNAMIC MODEL OF A HELICOPTER

The overall vehicle equations of motion are derived. The forces and moments from the different elements of a helicopter, such as main rotor, tail rotor, fuselage and empennage, are discussed in this paper. The helicopter has six degree of freedom in its motion and it has nine state variables in general, which are u,v,w the aircraft velocity components at centre of gravity, p,q,r the aircraft roll, pitch and yaw rates about body reference axes, and θ,ϕ,ψ the Euler angles. To derive the equations of the translational and rotational motions of a helicopter, the helicopter is assumed to be a rigid body referred to an axes system fixed at the centre of mass of the aircraft, so the axes move with time varying velocity components under the action of the applied forces. The Euler angles define the orientation of the fuselage with respect to earth axes system (Padfield, 1996). There are four control inputs, which are, longitudinal cyclic stick (η_{1s}), lateral cyclic stick (η_{1c}) , collective lever (η_c) and pedal input (η_p) which control the helicopter's motion through $X,Y,Z,L,M,and\ N$. So the system equations are as follows

$$\dot{u} = rv - qw - g\sin\theta + X/M_a \tag{1a}$$

$$\dot{v} = pw - ru + g\cos\theta\sin\phi + Y/M_a \tag{1b}$$

$$\dot{w} = qu - pv + g\cos\theta\cos\phi + Z/M_a \tag{1c}$$

$$\dot{p} = \frac{(I_{yy} - I_{zz})qr}{I_{xx}} + \frac{I_{xz}\left(pq(I_{xx} - I_{yy}) + rq + I_{yy} - I_{zz}\right) + I_{xx}^2 pq}{\left(I_{xx}^2 + I_{xx}I_{zz}\right)} + pq + \frac{I_{xz}(N - L)}{\left(I_{xx}^2 + I_{xx}I_{zz}\right)} + L \tag{2a}$$

$$\dot{q} = \frac{1}{I_{yy}}\left[(I_{zz} - I_{xx})pr - I_{xz}\left(p^2 - r^2\right)\right] + \frac{1}{I_{yy}}M \tag{2b}$$

$$\dot{r} = \frac{I_{xz}\left(rq - I_{yy} - I_{zz}\right) - I_{xz}^2 pq + pq\left(I_{xx} - I_{yy}\right)}{\left(I_{xz}^2 + I_{xx}I_{zz}\right)} + \frac{N - I_{xz}L}{\left(I_{xz}^2 + I_{xx}I_{zz}\right)} \tag{2c}$$

$$\dot{\phi} = p + q\sin\phi\tan\theta + r\cos\phi\tan\theta \tag{3a}$$

$$\dot{\theta} = q\cos\phi - r\sin\phi \tag{3b}$$

$$\dot{\psi} = q\sin\phi\sec\theta + r\cos\phi\sec\theta \tag{3c}$$

where M_a and g are the mass of the helicopter and acceleration due to gravity, I_{xx}, I_{yy}, I_{zz} are the moment of inertia of the helicopter about x, y and z axes, and I_{xz} the aircraft product of inertia. The model (1) ~ (3) can be considered as a cascade connection nonlinear system, that is, it has the following form:

$$\dot{x} = f_1(x, y) \tag{4}$$

$$\dot{y} = f_2(x, y) + G(y)\ U \tag{5}$$

The overall external forces X, Y and Z along x, y, z axes and moments L, M, N about x, y, z axes can be written as

$$\begin{aligned} X = {} & \frac{1}{2}\rho(\Omega R)^2\pi R^2 a_0 s\cos\gamma_s\frac{2C_x}{a_0 s} - \frac{1}{2}\rho(\Omega R)^2\pi R^2 a_0 s\sin\gamma_s\frac{2C_Z}{a_0 s} \\ & + \frac{1}{2}\rho(\Omega R)^2\pi R^2 a_0 s S_p\overline{V}_F^2 C_{XF}(\alpha_F) \end{aligned} \tag{6}$$

$$\begin{aligned} Y = {} & \frac{1}{2}\rho(\Omega R)^2\pi R^2 a_0 s\left(\frac{2C_Y}{a_0 s}\right) + \frac{1}{2}\rho(\Omega R)^2\overline{V}_{FN}^2 S_{FN} C_{YFN}(\beta_{FN}) \\ & + \frac{1}{2}\rho(\Omega_T R_T)^2 a_{0T} s_T(\pi R_T)^2\left(\frac{2C_{TT}}{a_0 s_T}\right)F_T \\ & + \frac{1}{2}\rho(\Omega R)^2 S_S\overline{V}_F^2 C_{YS}\frac{v_A}{V_F} \end{aligned} \tag{7}$$

$$\begin{aligned} Z = {} & \frac{1}{2}\rho(\Omega R)^2\pi R^2 a_0 s\left(\sin\gamma_s\left(\frac{2C_X}{a_0 s}\right) + \cos\gamma_s\left(\frac{2C_Z}{a_0 s}\right)\right) \\ & + \frac{1}{2}\rho(\Omega R)^2\overline{V}_T^2 S_{TP} C_{ZTP}(\alpha_{TP}) + S_p\overline{V}_F^2 C_{ZF}(\alpha_F) \end{aligned} \tag{8}$$

and

$$\begin{aligned} L = {} & \frac{-b}{2}K_\beta\beta_{1s} + h_R\left(\frac{1}{2}\rho\pi R^2(\Omega R)^2 a_0 s\left(\frac{2C_Y}{a_0 s}\right)\right) \\ & + h_T\cdot\frac{1}{2}\rho(\Omega_T R_T)^2 a_{0T} s_T(\pi R_T)^2\left(\frac{2C_{TT}}{a_{0T} s_T}\right)F_T \\ & + h_{FN}\cdot\frac{1}{2}\rho(\Omega R)^2\overline{V}_{FN}^2 S_{FN} C_{YFN}(\beta_{FN}) \end{aligned} \tag{9}$$

$$\begin{aligned} M = {} & \frac{-b}{2}K_\beta\beta_{1s} - h_R\cdot\frac{1}{2}\rho\pi R^2(\Omega R)^2 a_0 s\left(\cos\gamma_s\left(\frac{2C_X}{a_0 s}\right) - \sin\gamma_s\left(\frac{2C_Z}{a_0 s}\right)\right) \\ & + x_{cg}\cdot\frac{1}{2}\rho\pi R^2(\Omega R)^2 a_0 s\left(\sin\gamma_s\left(\frac{2C_X}{a_0 s}\right) + \cos\gamma_s\left(\frac{2C_Z}{a_0 s}\right)\right) \\ & + (l_{TP} + x_{cg})\frac{1}{2}\rho(\Omega R)^2\overline{V}_T^2 S_{TP} C_{ZTP}(\alpha_{TP}) \\ & + \frac{1}{2}\rho(\Omega R)^2 S_p l_F\overline{V}_F^2 C_{MF}(\alpha_F) \end{aligned} \tag{10}$$

$$\begin{aligned} N = {} & \frac{1}{2}\rho(\Omega R)^2\pi R^3 s a_0\left(\frac{2C_Q}{a_0 s} + \left(\frac{2I_R}{bI_\beta}\right)\frac{\Omega'}{\gamma}\right) \\ & - x_{cg}\left(\frac{1}{2}\rho\pi R^2(\Omega R)^2 a_0 s\left(\frac{2C_Y}{a_0 s}\right)\right) \\ & - (l_T + x_{cg})\frac{1}{2}\rho(\Omega_T R_T)^2 a_{0T} s_T(\pi R_T)^2\left(\frac{2C_{TT}}{a_0 s_T}\right)F_T \\ & - (l_{FN} + x_{cg})\frac{1}{2}(\Omega R)^2\overline{V}_{FN}^2 S_{FN} C_{YFN}(\beta_{FN}) \\ & + \frac{1}{2}\rho(\Omega R)^2 S_S l_F\overline{V}_F^2 C_{NF}(\beta_F) \end{aligned} \tag{11}$$

where ρ is air density, R and Ω are the main rotor blade radius and speed, a_0 and s the main rotor blade lift curve slope and solidity, and C_x, C_Y, C_z are the main rotor force coefficients in shaft axes. They are given by

$$\begin{bmatrix} C_X \\ C_Y \end{bmatrix} = \begin{bmatrix} \cos\Psi_w & -\sin\Psi_w \\ \sin\Psi_w & \cos\Psi_w \end{bmatrix}\begin{bmatrix} C_{XW} \\ C_{YW} \end{bmatrix} \tag{12}$$

$$\left(\frac{2C_Z}{a_0 s}\right) = -\left(\frac{2C_T}{a_0 s}\right) = -F_0^{(1)} \tag{13}$$

where Ψ_W is the side-slip angle and C_T is the main rotor thrust coefficient given by

$$C_T = \frac{T}{\rho(\Omega R)^2\pi R^2} \tag{14}$$

The main rotor force coefficients in the hub-wind axes C_{XW} and C_{YW} are can be obtained through equations (15) and (16) in terms of harmonic components of integrated blade aerodynamic loads and harmonics of flapping.

$$\left(\frac{2C_{XW}}{a_0 s}\right) = \left(\frac{F_0^{(1)}}{2} + \frac{F_{2c}^{(1)}}{4}\right)\beta_{1cw} + \frac{F_{1c}^{(1)}}{2}\beta_0 + \frac{F_{2s}^{(1)}}{4}\beta_{1sw} + \frac{F_{1s}^{(2)}}{2} \tag{15}$$

$$\left(\frac{2C_{YW}}{a_0 s}\right) = \left(\frac{F_0^{(1)}}{2} + \frac{F_{2c}^{(1)}}{4}\right)\beta_{1sw} - \frac{F_{1s}^{(1)}}{2}\beta_0 - \frac{F_{2s}^{(1)}}{4}\beta_{1cw} + \frac{F_{1c}^{(2)}}{2} \tag{16}$$

where β_0 is the coning angle and β_{1cw}, β_{1sw} are the first harmonic cyclic flapping angles. The harmonic components of integrated blade aerodynamic loads are given by the following expressions

$$F_0^{(1)} = \theta_0\left(\frac{1}{3} + \frac{\mu^2}{2}\right) + \frac{\mu}{2}\left(\theta_{1sw} + \frac{\overline{p}_w}{2}\right) + \left(\frac{\mu_z - \lambda_0}{2}\right) + \frac{1}{4}(1 + \mu^2)\theta_{tw} \tag{17}$$

$$F_{1s}^{(1)} = \frac{\alpha_{sw} + \theta_{1sw}}{3} + \mu\left(\theta_0 + \mu_z - \lambda_0 + \frac{2}{3}\theta_{tw}\right) \tag{18}$$

$$F_{1c}^{(1)} = \frac{\alpha_{cw} + \theta_{cw}}{3} - \frac{\mu\beta_0}{2} \tag{19}$$

$$F_{2s}^{(1)} = \frac{\mu}{2}\left\{\theta_{1cw} - \beta_{1sw} + \frac{\overline{q}_w - \lambda_{1cw}}{2} - \mu\beta_0\right\} \tag{20}$$

$$F_{2c}^{(1)} = -\frac{\mu}{2}\left\{\theta_{1sw} + \beta_{1cw} + \frac{\overline{p}_w - \lambda_{1sw}}{2} + \mu\left(\theta_0 + \frac{\theta_{tw}}{2}\right)\right\} \tag{21}$$

$$\begin{aligned} F_{1s}^{(2)} = {} & \frac{\mu^2}{2}\beta_0\beta_{1sw} + \left(\mu_z - \lambda_0 - \frac{\mu}{4}\beta_{1cw}\right)_{\alpha_{sw}} - \frac{\mu}{4}\beta_{1sw}\alpha_{cw} \\ & + \theta_0\left(\frac{\alpha_{sw}}{3} + \mu(\mu_z - \lambda_0) - \frac{\mu^2}{4}\beta_{1cw}\right) + \left(\frac{\alpha_{sw}}{4} + \frac{\mu}{2}\left(\mu_z - \lambda_0 - \frac{\beta_{1c}\mu}{4}\right)\right)\theta_{tw} \\ & + \theta_{1sw}\left(\frac{\mu_z - \lambda_0}{2} + \mu\left(\frac{3}{8}(\overline{p}_w - \lambda_{1sw}) + \frac{\beta_{1cw}}{4}\right)\right) \\ & + \frac{\mu}{4}\theta_{1cw}\left(\frac{\overline{q}_w - \lambda_{1cw}}{2} - \beta_{1sw} - \mu\beta_0\right) - \frac{\delta\mu}{a_0} \end{aligned} \tag{22}$$

$$F_{1c}^{(2)} = -2\beta_0\mu\left(\mu_z - \lambda_0 - \frac{4}{3}\mu\beta_{1cw}\right) + \left(\mu_z - \lambda_0 - \frac{3}{4}\beta_{1cw}\mu\right)\alpha_{cw} - \frac{\mu}{4}\beta_{1sw}\alpha_{sw} + \theta_0\left(\frac{\alpha_{cw}}{3} - \frac{\mu}{2}\left(\beta_0 + \frac{\mu}{2}\beta_{1sw}\right)\right) + \theta_{tw}\left(\frac{\alpha_{cw}}{4} - \mu\left(\frac{\beta_0}{3} + \frac{\mu}{8}\beta_{1sw}\right)\right) + \theta_{1cw}\left(\frac{\mu_z - \lambda_0}{2} + \frac{\mu}{4}\left(\frac{\bar{p}_w - \lambda_{1sw}}{2} - \beta_{1cw}\right)\right) + \frac{\mu}{4}\theta_{1sw}\left(\frac{\bar{q}_w - \lambda_{1cw}}{2} - \beta_{1sw} - \mu\beta_0\right) \quad (23)$$

where θ_0 is the main rotor collective pitch and is given by

$$\theta_0 = \frac{(g_{c0} + g_{c1}\eta_c) + (k_g\Delta n)}{1 + \tau_{c4}s} \quad (24)$$

In equation (24), g_{c0} and g_{c1} are the collective gearing constants, k_g and Δn the autostabiliser feed back gain and aircraft normal acceleration increment, and η_c the collective lever variable (control input).

$\theta_{1cw}, \theta_{1sw}$ blade cyclic pitch components in hub-wind axes are defined by

$$\begin{bmatrix}\theta_{1sw}\\ \theta_{1cw}\end{bmatrix} = \begin{bmatrix}\cos\Psi_w & \sin\Psi_w\\ -\sin\Psi_w & \cos\Psi_w\end{bmatrix}\begin{bmatrix}\theta_{1s}\\ \theta_{1c}\end{bmatrix}, \quad (25)$$

where θ_{1s}, θ_{1c} are the longitudinal and lateral cyclic pitch they are determined by

$$\theta_{1s} = \left[\frac{g_{1s0} + g_{1s1}\eta_{ls} + g_{sc0} + g_{sc1}\eta_c + k_\theta\theta + k_q q + k_{1s}(\eta_{ls} - \eta_{ls0})}{1 + \tau_{c1}s}\right]\cos\psi_F + \left[\frac{g_{1c0} + g_{1c1}\eta_{lc} + k_\phi\phi + k_p P + k_{1c}(\eta_{lc} - \eta_{lc0})}{1 + \tau_{c2}s}\right]\sin\psi_F \quad (26)$$

$$\theta_{1c} = \left[\frac{g_{1c0} + g_{1c1}\eta_{lc} + k_\phi\phi + k_p P + k_{1c}(\eta_{lc} - \eta_{lc0})}{1 + \tau_{c2}s}\right]\cos\psi_F - \left[\frac{g_{1s0} + g_{1s1}\eta_{ls} + g_{sc0} + g_{sc1}\eta_c + k_\theta\theta + k_q q + k_{1s}(\eta_{ls} - \eta_{ls0})}{1 + \tau_{c1}s}\right]\sin\psi_F \quad , \quad (27)$$

where k_ϕ, k_p, k_θ *and* k_q feedback gains, k_{1c} *and* k_{1s} are feed forward gains. η_{1s0}, *and* η_{1c0} are constants, adjustable by the pilot and η_{1c}, η_{1s} are lateral and longitudinal cyclic stick variables (control inputs).

The tail rotor provides control for the *yaw,* whose only responsibility is to provide a sideways thrust force and thereby produce a yawing moment about the main rotor shaft (Newman, 1994, Leishman, 2000) i.e. contributes the external force Y, moments L, and M (see equations (7), (9) and (11)). Tail rotor contributaion can be determined by

$$\left(\frac{2C_{TT}}{a_{0T}s_T}\right) = \frac{1}{3}\cdot\frac{\theta_{0T} + K_3\left(\frac{n_\beta}{\lambda_\beta^2}\right)_T\frac{4}{3}(\mu_{zT} - \lambda_{0T})}{1 - K_3\left(\frac{n_\beta}{\lambda_\beta^2}\right)_T(1 + \mu_T^2)}\cdot\left(1 + \frac{3}{2}\mu_T^2\right) + \left(\frac{\mu_{zT} - \lambda_{0T}}{2}\right), \quad (28)$$

where θ_{0T} is a tail rotor pitch (control input) and is given by

$$\theta_{0T} = \frac{g_{t0} + g_{t1}(g_{ct0}(1 - \eta_P) + (1 - 2g_{ct0})\eta_C) + k_w(\psi - \psi_H) + k_r r}{1 + \tau_{c3}s} \quad (29)$$

In equation (29), g_{t0}, g_{t1} are pedals gearing constants, and g_{ct0} is the pedal cable gearing constant. η_c, η_p are the collective liver variable and pedal variable, which are the control inputs.

Almost all of the performance characteristics of a helicopter depend on the power-plant performance (Prouty, 1986). Here a simplified model for a helicopter rotor-speed, associated engine and rotor governor dynamics formulae are presented as follows.

$$\dot{\Omega} = \frac{1}{I_R}(Q_E - Q_R - G_T Q_T) + \dot{r} \quad (30)$$

$$Q_E = -K_3(\Omega - \Omega_i) \quad (31)$$

$$Q_R = \frac{1}{2}\rho(\Omega R)^2\pi R^3 s a_0\left(\frac{2C_Q}{a_0 s}\right) \quad (32)$$

$$Q_T = \frac{1}{2}\rho(\Omega_T R_T)^2\pi R_T^3 s_T a_{0T}\left(\frac{2C_{QT}}{a_{0T}s_T}\right) \quad (33)$$

In time domain, differential equation can be written as

$$\ddot{Q}_E = \frac{1}{\tau_{e1}\tau_{e3}}\left\{-(\tau_{e1} + \tau_{e3})\dot{Q}_E - Q_E + K_3(\Omega - \Omega_i + \tau_{e2}\dot{\Omega})\right\} \quad (34)$$

where Q_E, Q_R and Q_T are the engine, main rotor, and tail rotor torques, respectively. G_T is the tail rotor gear ratio. I_R is the moment of inertia of the rotating system. Ω_i is the idling rotor speed and K_3 overall engine/rotor speed gain.

Some assumptions are made to produce a closed form of expressions to the helicopter motion, which are shown below. The tail rotor flapping is ignored. For the main rotor, flapping angles are assumed small and the overall fuselage acceleration and blade weight effects are neglected. Yaw rate and sideslip rate, are assumed small compared with rotor angular rate Ω in the kinematics of blade motion. Especially, basic assumptions regarding to the rotor blade aerodynamics are summarised as follows:

- A constant, two-dimensional, lift curve slope is assumed.
- Compressibility effects are ignored.
- Stall and reversed flow effects are ignored.
- The induced velocity distribution, normal to the rotor disc, includes linear longitudinal and lateral variations, the value at the centre satisfying simple momentum considerations.
- Couplings from blade pitch and lag dynamics into flapping motion are ignored.
- Quasi-steady flapping and coning are used in the derivation of the reaction forces and moments on the fuselage, i.e., the interaction of disc tilt modes with fuselage mode are neglected.

These assumptions make it possible to integrate the aerodynamic loading analytically and hence produce the closed form of expressions for the rotor forces and moments.

III. Simulation Study of UH-60 Black Hawk Helicopter Using Flightlab

A similar approach has been taken for modelling main rotor and tail rotor apart from that the tail rotor flapping has been ignored. In Flightlab, the tail rotor component implemented is based on a simplified theoretical method of determining the characteristics of a lifting rotor in forward flight (report No 716 National Advisory committee for Aeronautics by F.J. Bailey, Jr,) which is called bailey rotor. The co-ordinate system of the bailey rotor has X forward into the free-stream airflow and Z in the direction of thrust. Rotor thrust and torque are calculated as functions of the blade tip loss factor by making the similar assumptions as mentioned above. Bailey derived rotor thrust and torque by analytically integrating the air-loads over the rotor blade span and averaging them over the azimuth. For the bailey rotor, using a reasonable initial value for the tail rotor thrust T_T, the thrust coefficient C_{TT} can be calculated from momentum theory as

$$C_{TT} = \frac{T_T}{\rho(\Omega_T R_T)^2 \pi R_T^2 k_{bl}} \tag{35}$$

The only difference from mathematical model described in (28) is that the blockage effect, k_{bl}, is introduced due to fin consideration. Total inflow λ_0 and the induced velocity v_i can be calculated by

$$\lambda_0 = \frac{-\mu_z + \frac{1}{2} C_{TT}}{\sqrt{\left(\mu^2 + \frac{1}{2} C_{TT}\right)}} \tag{36}$$

$$v_i = \mu_z - \lambda_0 \tag{37}$$

Using these values, an iterative procedure is performed to determine the values of the total inflow. This is done by using the equation below derived by applying momentum theory.

$$v_i = \frac{a_{0T} s_T}{2} \left[\frac{\mu_z t_{3,1} + \theta_0 t_{3,2} + \theta_1 t_{3,3}}{2\sqrt{(\mu^2 + \lambda_0^2)} + \frac{a_{0T} s_T}{2} t_{3,1}} \right] \tag{38}$$

where θ_0 and θ_1 are the blade collective pitch at the root and tip, respectively. λ_0 is the total inflow across the rotor disk. The values $t_{3,1}$, $t_{3,2}$ and $t_{3,3}$ are computed by

$$t_{3,1} = \frac{1}{2} B^2 + \frac{1}{4} \mu^2 \tag{39}$$

$$t_{3,2} = \frac{1}{2} B^3 + \frac{1}{2} B \mu^2 \tag{40}$$

$$t_{3,3} = \frac{1}{4} B^4 + \frac{1}{4} B^2 \mu^2 \tag{41}$$

Once the values of $t_{3,1}$, $t_{3,2}$ and $t_{3,3}$ have been determined within a reasonable tolerance, the induced velocity is again calculated using the above equation. From equation (36) the thrust coefficient is then recalculated using

$$C_{TT} = 2 v_i \sqrt{(\mu^2 \lambda_0^2)} \tag{42}$$

The rotor thrust can then be calculated with the following equation:

$$T_T = \frac{\pi}{180} \rho k_{bl} \pi (\Omega_T R_T^2)^2 , \tag{43}$$

where the blockage effect, k_{bl}, due to a fin is used to modify the rotor thrust as a function of the velocity which is given by,

$$k_{bl} = (1 - b_{t1}) \frac{u_A^2}{v_{bl}^2} + b_{t1}, \quad \{u_A \le v_{bl}\} \tag{44}$$

$$k_{bl} = b_{t2}, \quad \{u_A \ge v_{bl}\} \tag{45}$$

in equations (44) and (45) the transition velocity, v_{bl} and the tail blockage constants, b_{t1} and b_{t2}, are specified by the users. This allows calculating the root collective pitch θ_0 by

$$\theta_0 = \frac{\pi}{180} \left[-T_T \delta_3 \tan \delta_3 \right] + \theta_{bias} + \theta_c , \tag{46}$$

where θ_c is the commanded root collective pitch, θ_{bias} is a preset collective pitch bias, T_T is the tail rotor thrust, and δ_3 is the hinge skew angle for pitch-flap coupling. In addition, for the main rotor, in Flightlab model, aerodynamic effect has been taken into account. In more details, the aerodynamic components are numeric components that allow the computation of airloads, inflow, and interference. Airloads are computed to give the motion of the attached structural component and inflow is computed based on the airloads. Additionally, interference between the aerodynamic components can be computed.

Fuselage and empennage components are implemented in Flightlab model, by using simple aerodynamics laws, in which the forces and moments from these elements are given by functions of incident and sideslip angle. In the process of modelling due to the complex flow field around helicopter fuselages and the interaction of the main rotor wake with the fuselage, some difficulties are caused to construct the forces and moments equations. So the direct results from wind tunnel test data gathered from various sources are used (Biggers, 1962, Wilson, *et at,* 1975). The engine output torque is controlled by the governor system that senses a change in rotor speed Ω and demands a fuel flow change ω_f. The fuel change is represented as a single lag.

$$\tau_{e1} \dot{\omega}_f + \omega_f = K_{e1} \Delta\Omega , \tag{47}$$

where τ_{e1} and K_{e1} are the time constant and gain respectively. K_{e1} is the slope of the droop in the rotor speed from flight idle to maximum contingency fuel flow.

$$\Delta\Omega = \Omega - \Omega_i , \tag{48}$$

where $\Delta\Omega$ and Ω_i are the changes in rotor speed and flight idle rotor speed, respectively. The engine torque Q_E response to the

fuel flow change is described by a lag responding to fuel flow and flow rate

$$\tau_{e3}\dot{Q}_E + Q_E = K_{e2}\left(\tau_{e2}\dot{\omega}_f + \omega_f\right) \quad , \tag{49}$$

where K_{e2} is the gain and τ_{e2} , τ_{e3} are the time constants. Combining the above results gives a second order ordinary differential equation as follows

$$\ddot{Q}_E = \frac{1}{\tau_{e1}\tau_{e3}}\left\{-\left(\tau_{e1}+\tau_{e3}\right)\dot{Q}_E - Q_E + K_3\left(\Omega - \Omega_i + \tau_2\dot{\Omega}\right)\right\} \tag{50}$$

The equation is further normalised by maximum engine torque $Q_{E\max}$ as

$$\left(\tau_{e1}\tau_{e3}\right)\ddot{\overline{Q}}_E + \left(\tau_{e1}+\tau_{e3}\right)\dot{\overline{Q}}_E + \overline{Q}_E = \frac{K_3}{Q_{E\max}}\left[\left(\Omega - \Omega_i\right) + \tau_{e2}\dot{\Omega}\right] \tag{51}$$

Where $$\overline{Q}_E = \frac{Q_E}{Q_{E\max}} K_3 = K_{e1}K_{e2} \tag{52}$$

With $$K_3 = -\frac{Q_{E\max}}{\Omega_i\left(1-\frac{\Omega_m}{\Omega_i}\right)} \tag{53}$$

Ω_m is the rotor speed at maximum contingency fuel flow. Similarly in the mathematical model the simplified free turbine engine equations are also presented (see equation (34)).

In the flight control system, essentially, signals from the cyclic stick, collective lever and yaw pedals are transmitted to the main and tail rotor blades. Inter-links between collective lever, main rotor cyclic, and tail rotor collective pitch are also incorporated in the mathematical model. These pilot generated signals are combined with error signals from the stabilisation and automatic flight control systems and passed through a first order lag. The autostabiliser transmits signals from rate and attitude gyros to produce feedback control of roll through lateral cyclic, pitch through longitudinal cyclic and yaw through tail rotor collective. Feed forward signals are also incorporated in the cyclic loops for compensation also normal acceleration is fed back into the main rotor collective channel to reduce adverse rotor pitching moments at high forward speeds. In Flightlab model the control components are designed as multi input/ mulit output, linear and nonlinear sub system.

Simulation results are presented in this paper. The appropriate parameters for the UH-60 helicopters are used in the simulation studies are presented in the appendix Table A.1~A.4. Comparison result shows that there is a general agreement between the flight test data and the Flightlab GRM model simulation results. The flight test data were generated from the tests conducted for the UH-60 helicopter under very calm wind condition at Navy crows landing, California in September 1992 (Fletcher, 1993, Fletcher, 1995).

Simulation with two dynamics manoeuvres (Hover and 80Kts) for the four control input has been carried out. The model response was computed using the actual flight measured control positions. Both the flight data and the simulation data ware plotted in the same scale, which enables an easier comparison of the variables of interest, such as translational velocities (u, v, w), rotational velocities (p, q , r), Euler angles (ϕ, θ, ψ) and body axes accelerations (a_x, a_y, a_z) . In this paper longitudinal stick input simulation results are choosen as an example, which are shown In Fig. 3.1(a) and Fig. 3.1(b). The pilot's longitudinal stick input was used to drive the model in hover condition and 80Kts forward flight speed.

There exists reasonably good correlation with the flight data response, however some discrepancies are evident in the pitch rate (q) . Initially it starts with a good agreement but tends to differ in the long term, which might be an indication of that some unstable factors in real flight vehicle have not been included into the mathematical model.

IV. Concluding Remarks

The paper describes modelling and simulation study of a helicopter system. The mathematical model for a helicopter has been developed for simulation study and control analysis.
For the simulation study in the paper a UH-60 like Fightlab GRM model has been used. The model responses are compared with UH-60 flight test data in both hover and 80Kts forward flight conditions. Correlation in the main is satisfactory but anomalies are present. The possible reasons for those anomalies are suggested. Overall satisfactory results are achieved. Simulation analyses with the mathematical model itself are currently undergoing and the results will be investigated for the analysis of the system stability and control.

References

[1] Biggers, J.C., McCloud, J.L., and Patterakis, P., Wind tunnel tests on two full scales helicopter fuselage, NASA TN-D-154-8, 1962.

[2] Bramwell, A.R.S., Helicopter Dynamics, Edward Arnold 1976

[3] Fletcher, J.W., Identification of UH-60 stability derivative models in hover from flight test data, AHS, May 1993.

[4] Fletcher, J.W., A model structure for identification of linear models of the UH-60 helicopter in hover and forward flight, NASA TM 110362, August 1995

[5] Hilbert, K.B., A Mathematical Model of the UH-60 Helicopter, NASA TM-85890, 1984.

[6] Leishman, G., The Principles of Helicopter Aerodynamics, Cambridge University Press 2000.

[7] Newman, S., The Foundations of Helicopter Flight, Arnold, A member of the Hodder Headline Group, London 1994

[8] Padfield, G.D., "Helicopter Dynamics and Flight Control, Blackwell Science Ltd., 1996.

[9] Prouty, R., Helicopter Performance, Stability and Control, PWS Publishers, 1986.

[10] W.Z. Stepniewski and C.N. Keys, "Rotary-Wing Aerodynamics", Dover Publications, Inc, New York, Vol. 2, 1984.

[11] J.C. Wilson and R.E. Mineck, "Wind Tunnel investigation of helicopter – Rotor wake effects on three helicopter fuselage model", NASA TM-X-3185, 1975

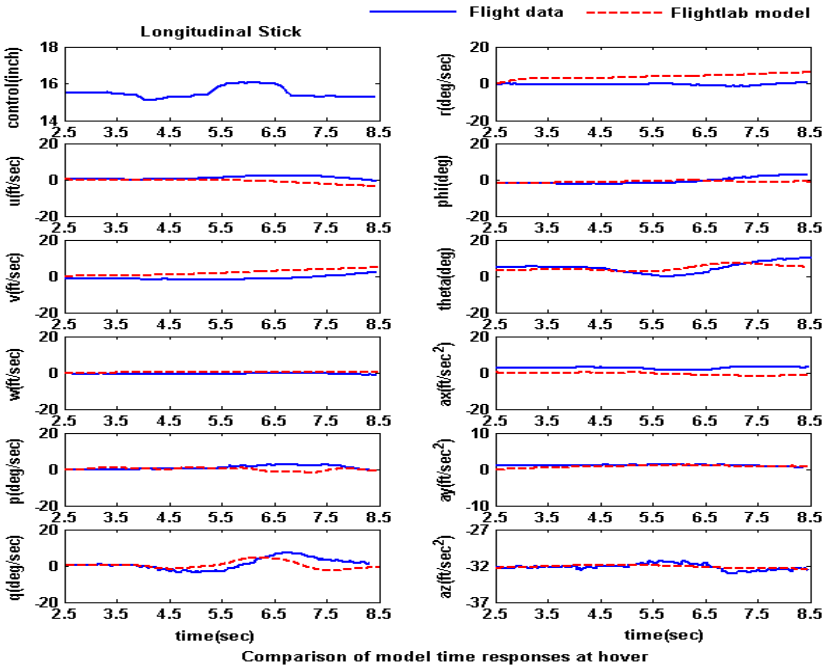


Figure. 3.1(a) Comparison of helicopter dynamic responses at hover between flight test data and FGR model for longitudinal stick input

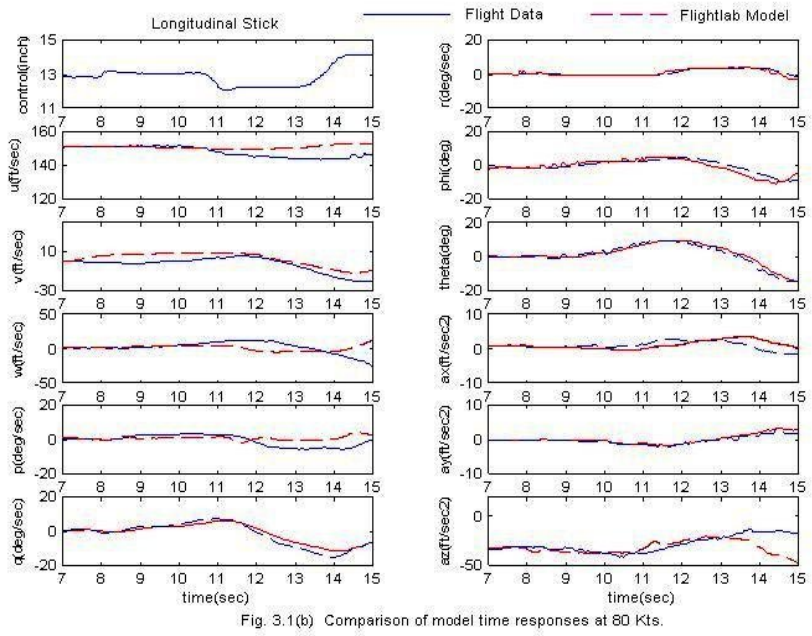


Figure. 3.1(b) Comparison of helicopter dynamic responses at 80Kts between flight test data and FGR model for longitudinal stick input

APPENDIX A

PARAMETER OF UH-60 HELICOPTER

UH-60 helicopter data are gathered from various sources are presented below (eg. Hilbert, 1984).

Table A.1 Aircraft mass and inertia:

Description	Symbol	UH-60 value	Units
mass of the helicopter	M_a	15350	lb
aircraft roll inertia	I_{xx}	5629	$slug\ ft^2$
aircraft pitch inertia	I_{yy}	40000	$slug\ ft^2$
aircraft yaw inertia	I_{zz}	37200	$slug\ ft^2$
aircraft product of inertia	I_{xz}	1670	$slug\ ft^2$
centre of gravity location	-	(36.0 0 4.7)	ft
Fuselage reference pt.	-	(34.6 0 23.4)	ft

Table A.2 Main rotor group:

Description	Symbol	UH-60 value	Units
main rotor speed	Ω	27.0	rad / sec
main rotor blade radius	R	26.83	ft
blade lift curve slope	a_0	5.73	rad^{-1}
main rotor solidity	s	0.08210	-
rotor shaft forward tilt	γ_s	0.05236	rad
rotor thrust coefficient	C_T	0.1846	-
number of blades	b	4	-
blade lock number	$\gamma(\gamma_0)$	8.1936	-
rotor inertia number	η_β	1.0242	-
flap frequency ratio	λ_β	1	-
linear blade twist	θ_{tw}	-0.3142	rad
Z co-ordinate of rotor hub	h_R	31.5	ft
mixing angle	ψ_F	0.175	rad
blade chord	c	1.73	ft
flapping spring const.	K_β	0	-
c.g. location fwd.of fuselage ref. Point	x_{cg}	1.4	ft
Stiffness number	S_β	0	-
blade profile drag coefficient	δ_0	-0.0216	-
air density	ρ	0.002473	$slug/ft^3$
blade flapping moment of inertia	I_β	3.10	$slug\ ft^2$

Table A.3 Empennage:

Description	Symbol	UH-60 value	Units
tail plane area	S_{TP}	45.0	ft
lift curve slope at zero incident	a_{0TP}	4	-
Location aft of fuselage reference point	l_{TP} , l_{FN}	70.0	ft
fin area	S_{FN}	32.3	ft^2

Table A.4 Tail rotor group:

Description	Symbol	UH-60 value	Units
tail rotor blade radius	R_T	5.5	ft
tail rotor speed	Ω_T	124.62	rad / s
blade lift curve slope	a_{0T}	5.73	rad^{-1}
tail rotor solidity	s_T	0.1875	-
fin blockage factor	F_T	-0.402	-
tail rotor inertia number	$(n_\beta)_T$	0.4223	-
flap frequency ratio	$(\lambda_\beta)_T$	1.0	-
tail rotor location aft of fuselage reference point	l_T	73.2	ft
Negative z co-ordinate of hub	h_T	32.5	ft
linear blade twist	θ_{tw}	-18.0	deg
Number of rotor blade	b	4	-
blade profile drag coefficient	δ_{0T}	-0.0216	-
blade lift dependent drag coefficient	δ_{2T}	0.40	-
pitch/flap coupling (δ_3)	k_3	0.700	-
blade lock number	$(\gamma)_T$	3.378	-

8 Robust Design of Terminal Iterative Learning Control with -synthesis Approach Applied for thermoforming oven control

Robust Design of Terminal Iterative Learning Control with μ-synthesis Approach

Applied for thermoforming oven control

Gauthier, Guy
Department of automated production engineering
École de technologie supérieure
1100 Notre-Dame Ouest, Montreal, Canada
guy.gauthier@etsmtl.ca

Boulet, Benoit
Department of electrical and computer engineering
McGill University
3480 University, Montreal, Canada
benoit.boulet@mcgill.ca

***Abstract*—This paper presents a robust design approach for the Terminal Iterative Learning Control (TILC) algorithm based on the μ-synthesis approach. TILC is used to control the reheat phase of plastic sheets in a thermoforming oven. TILC adjusts the heater temperature setpoints so that the plastic sheet temperature measured at the end of the reheat cycle converges to a desired temperature after a few cycles. Simulation results are included to show the effectiveness of this robust TILC algorithm.**

Keywords—Terminal Iterative Learning Control, robust control, μ-synthesis, thermoforming.

I. INTRODUCTION

In the thermoforming industry, the reheat phase is an important part of the process, since plastic sheets have to be heated to the right temperature before being molded [1-4]. The heater temperature setpoints are adjusted manually, by trial and error. This manual adjustment takes some cycles to complete, and results in monetary loss owing to the production of rejected parts.

This problem, observed in the plastics industry, has led to the idea of a cycle-to-cycle control approach to automatically tune the heater temperature setpoints [5]. To achieve this efficiently, the Terminal Iterative Learning Control (TILC) algorithm can be used to tune the setpoint temperature of heaters in a thermoforming oven [6-8]. The approach calls for the installation of temperature sensors measuring the surface temperature of the plastic sheet at the end of the cycle [4, 9].

TILC adjusts the heater temperature setpoints so that the sheet surface temperature converges to a desired temperature profile at the end of the heating cycle [10].

For the experimental oven, since it is small, the temperature sensors are inside the oven. But, for a bigger oven, the use of temperature scanner at the output of the oven is cheaper than using a lot a sensor inside the oven, so we need to use TILC.

This paper concerns the use of μ-synthesis as a tool to design TILC controllers. The μ-synthesis approach has been successfully used by other researchers for ILC – see [11, 12], and [13]. While robust ILC design has been studied extensively, robust TILC has attracted much less attention. TILC was introduced first in [14] and then in a PhD thesis in the same year [15]. High-order TILC, presented in [14-17], has been proposed to improve robustness. TILC is a variant of ILC, the main difference between them being that ILC has access to measurements sampled during the entire cycle, while TILC only has access to measurements sampled at the end of the cycle [14, 18].

Section II presents the system used to design TILC. Section II introduces the μ-synthesis concepts required to carry out the robust design, such as weighting functions and their parameters. Simulation results, using a TILC algorithm created with μ-synthesis, are shown in section IV. Section V concludes the paper.

II. PROBLEM SETUP

The system on which we apply the TILC algorithm is a thermoforming machine, but the algorithm can be applied to any system that behaves in a repetitive way. A linearized system is used to design TILC, and has been defined by [4, 6, 7, 10]:

$$\begin{aligned} \dot{x}_k(t) &= Ax_k(t) + Bu_k \\ y_k(t) &= Cx_k(t) \end{aligned} \tag{1}$$

In (1), $t \in \mathbb{R}$ and $k \in \mathbb{N}$ represent the time within the cycle and the cycle number respectively. The cycle length T has a fixed duration, because it is an important assumption of the TILC approach. Matrices A, B, and C are time-invariant. The state vector $x_k(t) \in \mathbb{R}^n$ expresses the temperature at n points on the plastic sheet. The input vector $u_k \in \mathbb{R}^m$ contains the temperature of the heaters, and those temperatures are maintained constant during the entire reheat cycle. The surface temperatures of the plastic sheet are in the output vector $y_k(t) \in \mathbb{R}^p$.

The control task is to update the control the input u_k, such that the sheet surface temperatures converge to a

desired terminal value vector $y_d \in \mathbb{R}^p$ at time T. For linear systems, the terminal output is

$$y_k(T) = \Gamma x_k(0) + \Psi u_k . \tag{2}$$

The matrix $\Gamma \in \mathbb{R}^{p\times n}$ is used to obtain the zero-input response, and is defined as

$$\Gamma = Ce^{AT} . \tag{3}$$

The matrix $\Psi \in \mathbb{R}^{p\times m}$ is used to obtain the zero-state response, and is defined by

$$\Psi = C\int_0^T e^{A(T-\tau)}Bd\tau . \tag{4}$$

To put the emphasis on the cycle domain, expressed by k, the notation can be changed, and so (2) is rewritten as

$$y_T[k] = \Gamma x_0[k] + \Psi u[k], \tag{5}$$

where $y_T[k] \coloneqq y_k(T)$, $u[k] \coloneqq u_k$, and $x_0[k] \coloneqq x_k(0)$.

Since it is a discretized system in the cycle domain, the z-transform is a useful tool for analyzing it. Then, the z-transform of (5) for the cycle domain is

$$\hat{y}_T(z) = \Gamma \hat{x}_0(z) + \Psi \hat{u}(z) \tag{6}$$

where the z-domain variables have a caret above them.

In the z-domain, the TILC algorithm is defined by

$$\hat{u}(z) = C(z)\left(\hat{y}_d(z) - \hat{y}_T(z)\right). \tag{7}$$

The closed-loop transfer function of the system (6) controlled with the TILC in (7) is expressed by

$$\hat{y}(z) = \left(I_p + \Psi C(z)\right)^{-1}\left\{\Psi C(z)\hat{y}_d(z) + \Gamma \hat{x}_0(z)\right\}, \tag{8}$$

or

$$\hat{u}(z) = C(z)\left(I_m + \Psi C(z)\right)^{-1}\left\{\hat{y}_d(z) - \Gamma \hat{x}_0(z)\right\}. \tag{9}$$

Lemma 1: The closed loop-system is internally stable if and only if the following matrix is invertible:

$$\begin{bmatrix} I_m & C(z) \\ -\Psi & I_p \end{bmatrix} \tag{10}$$

for all z outside the unit circle.

Proof: The proof can be found in [20, 21]. □

An example of a first-order TILC algorithm (since it is an integrator) would be

$$C(z) = (z-1)^{-1}\Psi^+ \tag{11}$$

where the + exponent represents the pseudoinverse operator.

Theorem 1: Suppose that a system represented by the matrix Ψ is controlled by the TILC algorithm expressed by (11). Then, the closed-loop system is internally stable.

Proof: From Lemma 1, the system is internally stable if (10) is invertible:

$$\begin{bmatrix} I_m & C(z) \\ -\Psi & I_p \end{bmatrix} = \begin{bmatrix} I_m & (z-1)^{-1}\Psi^+ \\ -\Psi & I_p \end{bmatrix}. \tag{12}$$

The determinant of (12) is

$$\begin{aligned} p(z) &\coloneqq \det\left(\begin{bmatrix} I_m & C(z) \\ -\Psi & I_p \end{bmatrix}\right) \\ &= \det\left(I_p + \Psi\Psi^+(z-1)^{-1}\right) \end{aligned} \tag{13}$$

Then,

$$p(z) = (z-1)^{-p}\det\left(I_p(z-1) + \Psi\Psi^+\right). \tag{14}$$

The roots of (14) ($p(z) = 0$) correspond to the poles of the closed-loop system. The zeros of $p(z)$ are also the eigenvalues of $I_p - \Psi\Psi^+$. There are two cases:

If $\Psi\Psi^+ = I_p$, then the eigenvalues are all equal to 0 and the system is stable.

If $\Psi\Psi^+ \neq I_p$, then the eigenvalues are equal to 0 or 1, since $I_p - \Psi\Psi^+$ is idempotent, like $\Psi\Psi^+$. The 1 eigenvalues are canceled out by the $(z-1)$ term in the denominator of (14), leaving the 0 eigenvalues, implying that the closed loop system is stable.

Then, in both cases, the closed-loop system is stable. □

The closed-loop system must remain stable, even when the parameters of the system are uncertain. However, fast convergence is needed to minimize the number of plastic sheets wasted.

III. The Mu-Synthesis Approach

The μ-synthesis approach can be used to design a robust TILC algorithm by tuning the filter parameters. Figure 1 shows the detailed block diagram of the system matrix with its uncertainties [8]. The system matrix Ψ is related to the nominal one, Ψ_0, with the weighting functions and matrices necessary to define the uncertainty of the system.

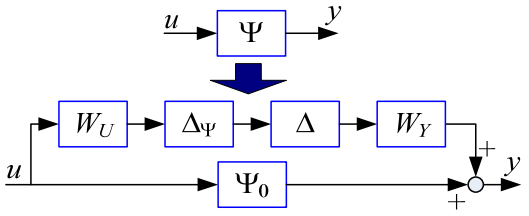


Figure 1: Representation of system uncertainty

We need to define each matrix appearing in this figure before proceeding.

The control system is expressed by its nominal matrix $\Psi_0 \in \mathbb{R}^{p\times m}$, and the uncertainty amplitudes on each entry of the system matrix Ψ are expressed by each entry $\Delta_{\Psi,ii} \in \mathbb{R}_+$, $\forall i \in \{1,2,\cdots,r\}$ of the real diagonal matrix

Δ_Ψ. The size of matrix Δ_Ψ depends on the number of uncertain real entries in the system, but the maximum size is $pm \times pm$.

The real diagonal matrix Δ is such that each entry on the main diagonal is strictly smaller than 1 ($|\Delta_{ii}| < 1$). Hence, the size of the matrix Δ is the same as Δ_Ψ. To associate each entry $\Psi_{0,ij}$ of Ψ_0 with the corresponding uncertainty amplitude $\Delta_{\Psi,kk}$, we need two real matrices identified by W_U and W_Y. Then, the uncertain system can be written as

$$\Psi = \Psi_0 + W_Y \Delta \Delta_\Psi W_U . \quad (15)$$

When all parameters of the system Ψ are uncertain, the matrix $W_U \in \mathbb{R}^{pm \times m}$ is

$$W_U = \begin{bmatrix} I_m & I_m & \cdots & I_m \end{bmatrix}^T \quad (16)$$

where the identity matrix I_m is repeated p times, and matrix $W_Y \in \mathbb{R}^{p \times pm}$ is

$$W_Y = I_p \otimes \begin{bmatrix} 1 & 1 & \cdots & 1 \end{bmatrix} \quad (17)$$

where the 1 is repeated m times.

When some parameters of the system have no uncertainty, the corresponding lines and columns have to be removed from the matrices, as shown in Figure 2. Hence, this operation reduces the size of all the matrices appearing in this figure [8].

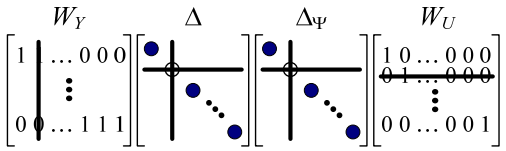


Figure 2: Removal of a 0 uncertainty

The uncertain system Ψ is connected to a cycle-to-cycle control to close the loop in the cycle domain. Figure 3 shows the complete block diagram of the system.

The TILC controller $C(z)$ results from combining the blocks $\tilde{C}(z)$ and $z^{-1}I$. The part of the controller to be designed using the μ-synthesis approach is $\tilde{C}(z)$.

Since the robustness of the controller is obtained by the uncertainties expressed earlier, the performance specifications are included in the $W_1(z)$ matrix containing the weighting function of the main diagonal. Each entry of $W_1(z)$ has the following transfer function:

$$W_{1,ii} = \frac{1}{M_{1,i}} \left\{ \frac{\left(2 + M_{1,i}\omega_{1,i}\right)z - \left(2 - M_{1,i}\omega_{1,i}\right)}{\left(2 + \varepsilon_{1,i}\omega_{1,i}\right)z - \left(2 - \varepsilon_{1,i}\omega_{1,i}\right)} \right\} \quad (18)$$

where $i \in \{1, 2, \cdots, p\}$.

The parameters of the weighting function are [8]:

- $M_{1,i}$: the high-frequency gain of the inverse of $W_{1,ii}$, its purpose being to limit the high-frequency gain of the sensitivity function of the closed-loop system;
- $\varepsilon_{1,i}$: the low-frequency gain of the inverse of $W_{1,ii}$ – the steady state error of the closed-loop system will remain under this value;
- $\omega_{1,i}$: the frequency where the gain of the inverse is equal to 1, which will determine the speed of convergence of the closed-loop system.

The block diagram shown in Figure 3 can be reorganized into the one shown in Figure 4. The TILC and uncertainty are put into distinct blocks, and all the other blocks are grouped into a block named *N*. Figure 4 can be simplified to an $N\Delta\tilde{C}$ representation, as shown in Figure 5.

From the block diagram in Figure 5, we can write

$$N := \begin{bmatrix} 0 & 0 & \Delta_\Psi W_U z^{-1} \\ -W_1 W_Y & W_1 & -W_1 \Psi_0 z^{-1} \\ -W_Y & I & -\Psi_0 z^{-1} \end{bmatrix} . \quad (19)$$

This *N* matrix is the main component of the following relationship between input and output shown in Figure 5:

$$\begin{bmatrix} p \\ z \\ e \end{bmatrix} = N \begin{bmatrix} q \\ y_d \\ u' \end{bmatrix} . \quad (20)$$

The μ-synthesis consists of finding a controller $\tilde{C}(z)$ that minimizes the maximum value of μ. The objective function to minimize is given by [8, 19]

$$\tilde{C}(z) = \arg \min_{\text{stabilizing } C(z)} \mu\left(\mathcal{F}_l(N(z), \tilde{C}(z)\right) . \quad (21)$$

In this function, the lower linear fractional transformation (LFT) is defined as

$$P := \mathcal{F}_l(N, \tilde{C}) = N_{11} + N_{12}\tilde{C}\left(I - N_{22}\tilde{C}\right)^{-1} N_{21} . \quad (22)$$

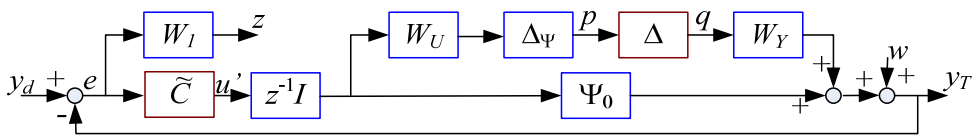


Figure 3: Block diagram of the closed-loop system with weighting functions

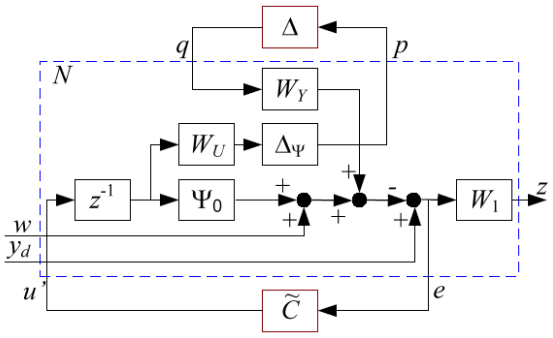


Figure 4: Reorganized block diagram

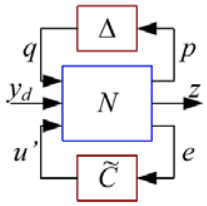


Figure 5: Simplified block diagram for μ-synthesis

In (22), N is partitioned as follows:

$$N_{11} = \begin{bmatrix} 0 & 0 \\ -W_1 W_Y & W_1 \end{bmatrix}, \tag{23}$$

$$N_{12} = \begin{bmatrix} \Delta_\Psi W_U z^{-1} \\ -W_1 \Psi_0 z^{-1} \end{bmatrix}, \tag{24}$$

$$N_{21} = \begin{bmatrix} -W_Y & I \end{bmatrix}, \tag{25}$$

$$N_{22} = -\Psi_0 z^{-1}. \tag{26}$$

The performance specification is expressed by a fictitious uncertainty matrix $\Delta_P \in \mathbb{C}^{p \times p}$. This matrix connects the output z of the weighting function W_1 to the input y_d.

The matrix Δ_P is such that

$$\|\Delta_P\| = \bar{\sigma}(\Delta_P) < 1 \tag{27}$$

at all frequencies. The matrix Δ_P is a full matrix of complex values. Both uncertainty matrices, Δ and Δ_P, are combined into one [19, 22]:

$$\mathbf{\Delta} = \begin{bmatrix} \Delta & 0 \\ 0 & \Delta_P \end{bmatrix}. \tag{28}$$

Note that $\mathbf{\Delta}$ satisfies

$$\|\mathbf{\Delta}\| = \bar{\sigma}(\mathbf{\Delta}) = \max\left(\bar{\sigma}(\Delta), \bar{\sigma}(\Delta_P)\right) < 1 \tag{29}$$

at all frequencies.

Since the norm of $\mathbf{\Delta}$ is smaller than 1, by the small gain theorem, the combination of the system expressed by matrix N and the controller $\tilde{C}$ must give a matrix P having a gain (or norm) smaller than 1 at all frequencies.

The μ-synthesis is a way to obtain the TILC controller $C = z^{-1}\tilde{C}$ for the system. The algorithm consists of finding a controller that minimizes the maximum value of μ. If the maximum value of μ is less than 1, controller C ensures a system with robust performance. If not, the closed-loop system violates at least one of the specifications, robust performance or robust stability.

To use the available software tools to perform μ-synthesis, the system must be converted from discrete to continuous time using the Tustin approach [8]. The controller found must then be converted back to discrete time. For this design, we use the "dkitgui" function in Matlab®.

IV. Simulation Results

The simulation is performed with a nonlinear model of the AAA thermoforming oven (this model is explained in [4, 6, 8]). The oven has 12 heater banks and 14 infrared temperature sensors (IRT) to measure the temperature at the surface of the plastic sheet. Figure 6 shows the location of the IRT sensors and the heaters in the oven. The TILC algorithm created with the μ-synthesis approach is designed with a linearized model of the thermoforming oven with the default parameters shown in Table 1. The weighting function parameters of W_1 are $M_1 = 2$, $\varepsilon_1 = 0.01$, and $\omega_1 = 0.5$. Those parameters will ensure a fast and monotonic convergence of the closed-loop system.

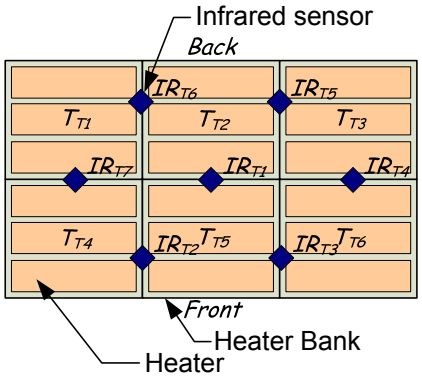


Figure 6: Heater and sensor location (bottom heaters and sensors at the same location, with subscript B)

TABLE I. Parameters used in the simulations

Parameter	Units	Default	Modified
Density	kg/m^3	950	1045
Specific heat	J/(kg·K)	1838	2021.8
Effective emissivity		0.45	0.495
Absortivity		300	350
Heat conduction	W/(m·K)	0.4	0.3
Convection factor	W/(m^2·K)	6	10

All simulations are performed on the nonlinear system, some with the default parameters and some with the modified parameters, to test the robustness of the closed-loop system.

A. *Two heater/two sensor configuration*

The first three simulations are performed with a two-heater, two-sensor configuration. All heaters above the plastic sheet are grouped together, as are all the bottom heaters. The desired terminal values are 150°C at IR_{T1} and 151°C at IR_{B1}. From the linearized model, the μ-synthesis-designed TILC algorithm is (with $\mu = 0.910$)

$$C(z) = (z - 0.995)^{-1} \begin{bmatrix} 1.8044 & -0.7342 \\ -0.7343 & 1.9061 \end{bmatrix}. \quad (30)$$

Using this TILC algorithm on the nonlinear model of the thermoforming machine, we obtain the maximum surface temperature error (MSTE) plot in Figure 7 (□). The temperature error falls below 5°C in 3 cycles.

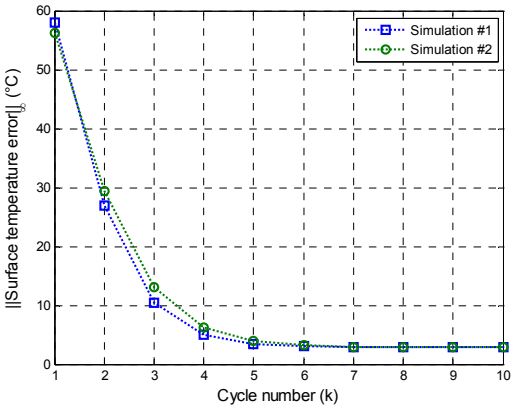


Figure 7: Surface temperature error, sim. #1 and #2

The initial heater temperature setpoints were set to 250°C in all simulations. Although a thermoforming oven operator might consider this initial adjustment to be unwise, that setting may make it possible to demonstrate the ability of the TILC to converge rapidly to the desired surface temperature profile.

With a non-linear model having the modified parameters, convergence is slower, since the maximum temperature error falls below 5°C in 6 cycles (○ in Figure 7). The energy transfer from the heater to the plastic sheet in this case seems to be less efficient, which is why the heater temperature setpoints are higher in simulation 2.

In the third simulation, the $(z - 0.995)^{-1}$ term in (30) is replaced by $(z - 1)^{-1}$. Then, the maximum surface temperature error converges to 0 and no steady-state error remains. However, the heater temperature setpoints are about 10°C higher.

B. *Six heater/sir sensor configuration*

For the fourth simulation, we consider an oven configuration with six heaters and six sensors. The TILC algorithm obtained gives the results shown in Figure 8. The convergence is slower and the maximum temperature error falls below 5°C in 8 cycles. For this algorithm, we obtained $\mu = 0.954$, with the weighting function parameter $\omega_1 = 0.125$.

C. *Ten heater/ten sensor configuration*

Finally, the last four simulations were performed with a 10 heater/10 sensor configuration. Using $\omega_1 = 0.125$ for the TILC design, the resulting TILC algorithm has a maximal μ equal to 0.997. The denominator of the controller obtained is $(z - 0.9988)$. The fifth simulation was performed with the system having the default parameters. Figure 9 shows a convergence to a value under 5°C in 7 cycles (□). A steady-state error remains.

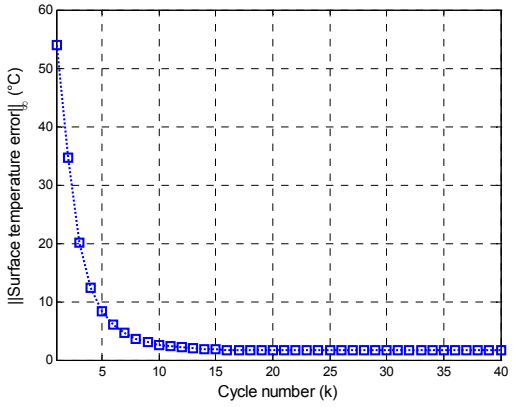


Figure 8: Surface temperature error, sim. #4

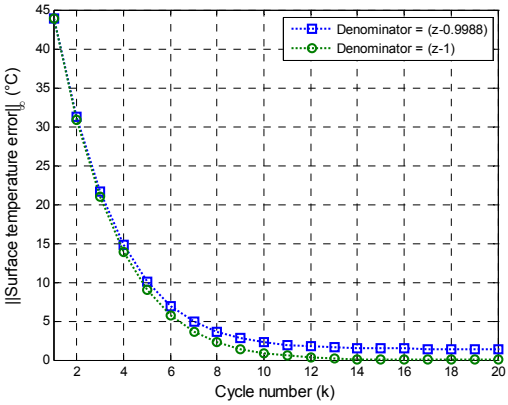


Figure 9: Surface temperature error, sim. #5 & #6

The measured surface temperatures have a monotonic convergence all measured temperatures.

For the sixth simulation, the denominator of the controller was changed to $(z - 1)$. With this new TILC, convergence is a little bit faster (○ in Figure 9) and the MSTE falls to 0.

The next simulation was performed with a system subject to a measurement noise having a standard deviation of 1°C, an ambient temperature drift of 1°C per cycle and slow sinusoidal variation of initial temperature. With the system having the default parameters, the MTSE converges below 5°C at the 8th cycle (Figure 10). Due to noise, drift, and initial temperature variation, the error does not fall to 0, even if the denominator is $(z - 1)$. From the 8th cycle, the MTSE remains under 5°C.

For the last simulation, with the modified parameters, the MTSE falls below 5°C at the 11th cycle. The effect of noise,

drift, and initial temperature variation is apparent on the heater temperature setpoints in Figure 11. The TILC algorithm is able to sustain slow variations of the oven temperature and variation of the initial temperature of the plastic sheet.

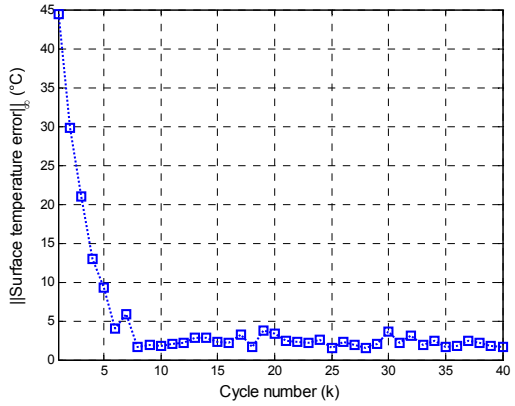


Figure 10: Surface temperature error, sim. #7

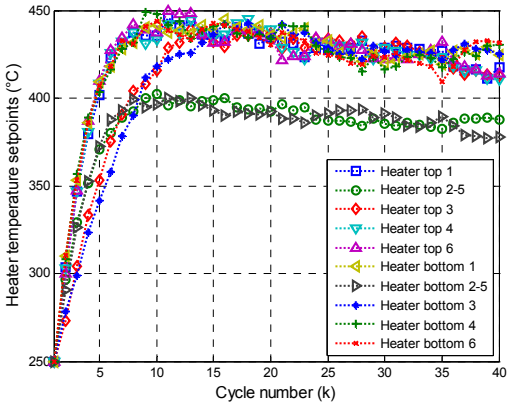


Figure 11: Heater temperature setpoints, sim. #8

V. CONCLUSION

The TILC algorithm design with μ-synthesis is able to produce a robust controller, as shown by the simulations. Now, there is a compromise to be made between robustness and performance. Each cycle with a maximum terminal surface temperature error over 10°C can lead to a wasted plastic sheet (for HDPE).

Then, the speed of convergence becomes an important specification defined by tuning the parameter ω_1 of the weighting function W_1.

Ideally, convergence must be monotonic, since an excessive overshoot on the heater temperature setpoints can lead to an overheated plastic sheet. The plastic can become too fluid and fall onto the bottom heaters, damaging the thermoforming oven. Therefore, the system must be robust in the face of slow parametric and environmental changes. μ-synthesis makes it possible to tune this compromise, so that the thermoforming oven will behave relatively well during the reheat phase, even with an unwise choice of initial heater temperature setpoints.

REFERENCES

[1] J. L. Throne, *Technology of thermoforming* (Hanser Publishers, 1996).

[2] J. L. Throne, *Thermoforming* (Hanser Publishers, 1987).

[3] B. Moore, *In-cycle control of the thermoforming reheat process*, Master's thesis, McGill University, Montreal, 2002.

[4] M. Ajersch, *Modelling and real-time control of sheet reheat phase in thermoforming*, Master's thesis, McGill University, Montreal, 2004.

[5] F. M. Duarte and J. A. Covas, Heating thermoplastic sheet for thermoforming solution to direct and inverse problems, *Plastic, Rubber and composite, 26*, 1997, 213-221.

[6] G. Gauthier, M. Ajersch, B. Boulet, A. Haurani, P. Girard, and R. DiRaddo, A new absorption based model for a sheet reheat in thermoforming, *SPE 2005 Annual Technical Conference*, Boston, MA, 2005.

[7] G. Gauthier and B. Boulet, Terminal iterative learning control applied to thermoforming reheat phase, *2006 International Symposium on Industrial Electronics*, Montreal, Canada, 2006, 353-357.

[8] G. Gauthier, *Terminal iterative learning for cycle-to-cycle control of industrial processes*, Ph.D. thesis, McGill University, Montreal, 2008.

[9] P. Girard, R. DiRaddo, V. Thompson, and B. Boulet, Advanced in-cycle and cycle-to-cycle on-line adaptive control for thermoforming of large thermoplastic sheets, *SAE Advances in Plastic Components, Processes and Technologies 2005*.

[10] G. Gauthier and B. Boulet, Robust design of terminal ILC with H∞ mixed sensitivity approach for thermoforming oven, *Fifth International Conference on Industrial Automation*, Montreal, Canada, 2007.

[11] Y.-J. Liang and D. P. Looze, Performance and Robustness Issues in Iterative Learning Control, *Proceedings of the 32nd Conference on Decision and Control*, San Antonio, Texas, 1993, 1990-1995.

[12] A. Tayebi and M. B. Mazemba, Robust Iterative Learning Control Design via μ-Synthesis, *Proceedings of the 2005 IEEE Conference on Control Applications*, Toronto, Canada, 2005.

[13] J.-H. Moon, T.-Y. Doh, and M. J. Chung, A Robust Approach to Iterative Learning Control Design for Uncertain Systems, *Automatica, 34(8)*, 1998, 1001-1004.

[14] Y. Chen, J.-X. Xu, and C. Wen, A High-Order Terminal Iterative Learning Control Scheme, *Proceedings of the 36th IEEE Conference on Decision and Control*, San Diego, California, 1997, 3771-3772.

[15] Y. Chen, *High-Order Terminal Iterative Learning Control: Convergence, Robustness and Applications*, Ph.D. thesis, Nanyang Technological University, Singapore, 1997.

[16] Y. Chen and J.-X. Xu, High-Order Terminal Iterative Learning Control with an Application to a Rapid Thermal Process for Chemical Vapor Deposition, in *Iterative Learning Control – Convergence, Robustness and Applications*, (Springer Verlag, 1999).

[17] Y. Chen, J.-X. Xu, T. H. Lee, and S. Yamamoto, An Iterative Learning in Rapid Thermal Processing, *Proceedings of the IASTED International Conference on Modeling, Simulation and Optimization,* Calgary, Canada, 1997.

[18] K. L. Moore, Iterative Learning Control: An Expository Overview, *Applied and Computational Control, Signals, and Circuits*, B.N. Datta, Editor, Birkhauser, 1999, 151-214.

[19] S. Toffner-Clausen, P. Anderson, and J. Soustrup, *Robust Control*, (Department of Control Engineering, Aalborg University, Aalborg, 2001).

[20] K. Zhou, *Essential of Robust Control* (Prentice Hall, Upper Saddle River, N.J., 1998).

[21] K. Zhou, J.C. Doyle, and K. Glover, *Robust and Optimal Control* (Prentice Hall, Upper Saddle River, N.J., 1996).

[22] D. W. Gu, P. H. Petkov, and M. M. Konstantivov, *Robust control design with MATLAB* (Springer-Verlag, New York, 2005).

9 Gain-scheduled H-inf control of a robotic manipulator with nonlinear joint friction

Gain-scheduled $\mathcal{H}_\infty$ control of a robotic manipulator with nonlinear joint friction

Seyed Mahdi Hashemi and Herbert Werner

Abstract— **This paper presents the LPV modelling and control of a robotic manipulator with a nonlinear joint friction model. A nonlinear dynamic model of the manipulator including viscous and Coulomb friction terms is obtained and the signum function in the friction model is approximated by a hyperbolic function in order to smooth such hard nonlinearity. A quasi-LPV model is derived and since it has a large number of affine scheduling parameters and a large overbounding, parameter set mapping is used to reduce conservatism and complexity in controller design by finding tighter parameter regions with fewer scheduling parameters. Then, a polytopic LPV gain-scheduled controller is synthesized and implemented experimentally on an industrial robot for a trajectory tracking task. The experimental results illustrate that the designed LPV controller outperforms a similar LPV controller based on a linear friction model, a model-based inverse dynamics and a decentralized PD controller in terms of tracking performance.**

I. Introduction

Modelling and control of systems in which friction occurs has been an active field of research during the last decades [1], [2] and [3]. A number of models have been developed for sliding and pre-sliding friction regimes [4] [5] and [6] which can capture various aspects of this phenomena. However, nonlinearities in most of them is a major challenge to design high performance controllers.

Linear parameter-varying (LPV) gain-scheduling techniques have evolved into a promising and effective framework for modern control applications. Their attractiveness lies in the extension of well-known linear optimal $\mathcal{H}_\infty$ control methods and the use of Linear Matrix Inequalities (LMIs), to the solution of nonlinear control problems, see e.g. [7], [8] and [9]. Many nonlinear systems can be converted into a quasi-LPV form, where the scheduling parameters may include system inputs, states, outputs and external signals. However, the number of reported successful implementation of LPV controllers on practical applications is still limited. Two major limiting problems in this regard are conservatism, among other things due to overbounding, and the large number of scheduling parameters [10].

Obviously, simply neglecting some dynamic terms of the plant model or heuristically freezing some scheduling parameters will in general not be an appropriate solution. Parameter set mapping based on principle component analysis (PCA) proposed in [10] helps to obtain LPV models with tighter parameter sets that have less overbounding. In addition, correlation between scheduling parameters is detected and insignificant directions in the parameter space can be neglected without loosing much information about the plant.

Using LPV gain-scheduling techniques, high performance controllers can be designed and implemented for nonlinear robotic manipulators. In [11], LPV modelling and control of a simple drive system with a nonlinear friction model is reported, where the scheduling parameters include some discontinuities which can be a challenge in many applications. In [12], [13] and [14], LPV gain-scheduling controllers were designed for robot models ignoring friction. In [15] and [16], LPV modelling and control of a robotic manipulator were reported, where a linear friction model was adopted.

This paper presents the application of LPV gain-scheduling techniques to modelling and identification of a two-degrees-of-freedom (2-DOF) robotic manipulator with a nonlinear friction model. Dynamic model of the robot including Coulomb and viscous friction terms is identified to derive a quasi-LPV model. Although there is no upper bound on the rate of change for scheduling parameters in polytopic LPV controller synthesis with a fixed Lyapunov function [7], existence of sudden changes with an infinite slope due to the Coulomb friction term in scheduling parameters can degrade the experimental results. Thus, a continuous hyperbolic function is used to smooth such hard nonlinearity.

A quasi-LPV model in polytopic form is constructed such that the state matrices depend on the scheduling parameters in an affine manner. Such a model has a large number of scheduling parameters, which makes the LPV controller synthesis conservative and computationally expensive. Thus, parameter set mapping is applied and less significant components are neglected and the accuracy of the approximated model is assessed. Afterwards, the model is discretized and a discrete-time polytopic LPV controller with a fixed Lyapunov function is designed and implemented on the robot for a trajectory tracking task. The experimental results illustrate that the designed LPV controller outperforms a similar LPV controller based on a linear viscous friction model, a model-based inverse dynamics and a decentralized PD controller in terms of tracking performance.

The contribution of this paper is that an smoothed nonlinear friction model is included in an LPV framework, which leads to an improvement in modelling accuracy and control performance. In addition, no ad-hoc simplification in the manipulator dynamics is made, but a systematic approach is used to derive a suitable LPV model for controller synthesis.

This paper is organized as follows. Dynamic and friction modelling of the robot is presented in section II. In Section

S. M. Hashemi and H. Werner are with the Institute of Control Systems, Hamburg University of Technology, Eissendorfer Str. 40, 21073 Hamburg, Germany, {seyed.hashemi, h.werner}@tu-harburg.de

III, LPV modelling and application of parameter set mapping for parameter reduction is discussed. The LPV controller synthesis is described in Section IV. Controller implementation and experimental results are given in Section V. The last section gives the conclusions.

II. DYNAMIC MODEL

The CRS A465 robotic manipulator shown in Fig. 1 has six rotational DOFs. In this paper, only the second and third joints of this robot, referred to as joint 1 and joint 2 respectively, are modeled and the other links are considered as parts of them, and are fixed during the experiments as shown in Fig. 2. These two links which represent the shoulder and elbow respectively, are the most challenging to control, since they are affected by gravity as well as inertial, centripetal, Coriolis and friction torques.

Fig. 1. The CRS A465 robot

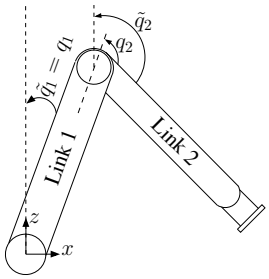


Fig. 2. Side view of the 2-DOF robot model

Using the Euler-Lagrange formulation, a rigid-body dynamic model of the 2-DOF manipulator is obtained as

$$\boldsymbol{\tau}(t) = \mathbf{D}(\mathbf{q}(t))\ddot{\mathbf{q}}(t) + \mathbf{c}(\mathbf{q}(t), \dot{\mathbf{q}}(t)) + \mathbf{g}(\mathbf{q}(t)) + \boldsymbol{\tau}_f(\dot{\mathbf{q}}(t)), \tag{1}$$

where $\boldsymbol{\tau}$ is the vector of joint torques (control inputs), and $\mathbf{q}$, $\dot{\mathbf{q}}$ and $\ddot{\mathbf{q}}$ are the vectors of joint positions, velocities and accelerations respectively, all belonging to $\mathbb{R}^2$. $\mathbf{D}$ is the inertia matrix and $\mathbf{c}$, $\mathbf{g}$ and $\boldsymbol{\tau}_f$ are the vectors of Coriolis-centrifugal, gravity and joint friction torques.

It is reported in [17] and [18] that a friction model containing Coulomb and viscous terms is appropriate for a wide range of robotic applications. This model is adopted here since Coulomb and viscous terms have a higher effect in comparison with the Stribeck term in this plant [19]

$$\tau_{f_i}(t) = f_{c_i}\text{sgn}(\dot{q}_i(t)) + f_{v_i}\dot{q}_i(t), \quad i = 1, 2, \tag{2}$$

where f_{c_i} and f_{v_i} are the Coulomb and viscous friction coefficients and sgn represents the signum function. Since the plant is not equipped with high resolution proximity sensors, friction models including the presliding regime can not be selected.

If joint positions are measured from a fixed coordinate axis shown in Fig. 2, some undesirable nonlinear terms in the dynamic equations will disappear, and developing an LPV model which depends affinely on scheduling parameters will become easier. Such a measurement is not possible in practice, but one can introduce new joint position variables $\tilde{q}_i$ with a fixed coordinate reference

$$\tilde{q}_1 = q_1, \tilde{q}_2 = q_1 + q_2. \tag{3}$$

Using (3), the dynamic model (1) is rewritten as

$$\boldsymbol{\tau}(t) = \tilde{\mathbf{D}}(\tilde{\mathbf{q}}(t))\ddot{\tilde{\mathbf{q}}}(t) + \tilde{\mathbf{c}}(\tilde{\mathbf{q}}(t), \dot{\tilde{\mathbf{q}}}(t)) + \tilde{\mathbf{g}}(\tilde{\mathbf{q}}(t)) + \tilde{\boldsymbol{\tau}}_f(\dot{\tilde{\mathbf{q}}}(t)), \tag{4}$$

$$\begin{aligned}
\tilde{\mathbf{D}} &= \begin{bmatrix} b_1 + b_3\cos(\tilde{q}_2 - \tilde{q}_1) & b_2 + b_3\cos(\tilde{q}_2 - \tilde{q}_1) \\ b_3\cos(\tilde{q}_2 - \tilde{q}_1) - b_8 & b_7 \end{bmatrix}, \\
\tilde{\mathbf{c}} &= \begin{bmatrix} b_3\dot{\tilde{q}}_1^2\sin(\tilde{q}_2 - \tilde{q}_1) - b_3\dot{\tilde{q}}_2^2\sin(\tilde{q}_2 - \tilde{q}_1) \\ b_3\dot{\tilde{q}}_1^2\sin(\tilde{q}_2 - \tilde{q}_1) \end{bmatrix}, \\
\tilde{\mathbf{g}} &= \begin{bmatrix} -b_4\sin(\tilde{q}_2) - b_5\sin(\tilde{q}_1) \\ -b_4\sin(\tilde{q}_2) \end{bmatrix}, \\
\tilde{\boldsymbol{\tau}}_f &= \begin{bmatrix} b_{10}\text{sgn}(\dot{\tilde{q}}_1) + b_6\dot{\tilde{q}}_1 \\ b_{11}\text{sgn}(\dot{\tilde{q}}_2 - \dot{\tilde{q}}_1) + b_{10}(\dot{\tilde{q}}_2 - \dot{\tilde{q}}_1) \end{bmatrix},
\end{aligned} \tag{5}$$

and the coefficients b_i, $i = 1 \ldots 11$ are linear combinations of dynamic and kinematic parameters, given in Appendix A.

Some scheduling parameters of the LPV model will include the signum function in (2), that may degrade the performance in experiments. Thus, it is approximated by a smooth hyperbolic tangent function shown in Fig. 3

$$\text{sgn}(\dot{q}_i(t)) \approx \tanh(\beta\dot{q}_i(t)), i = 1, 2, \tag{6}$$

where $\beta = 20$ is a smoothing coefficient that is chosen to make the scheduling trajectories smooth enough while providing suitable model accuracy. The unknown dynamic

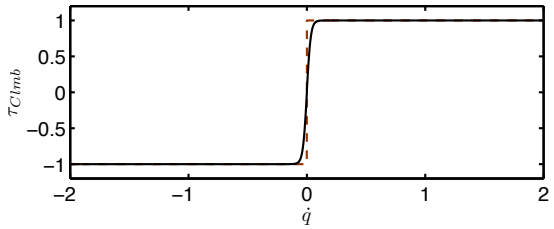


Fig. 3. Signum (dashed) and hyperbolic tangent functions(solid)

and friction parameters b_is are estimated using the linear least squares method, see [15] for more details.

The estimated parameters are given in Table I. To validate the accuracy of the estimated model, measured and predicted torques by the model for a cross validation trajectory are compared and plotted in Fig. 4, which illustrates a high

TABLE I

ESTIMATED INERTIAL AND FRICTION PARAMETERS

(WITH NON-SI UNITS)

b_1	0.0877	b_5	0.0407	b_9	0.5860
b_2	0.0241	b_6	0.4524	b_{10}	0.1603
b_3	-0.0075	b_7	-0.00957	b_{11}	0.7060
b_4	0.0154	b_8	0.0675		

accuracy. Having an accurate nonlinear model is of great importance in deriving a useful quasi-LPV model.

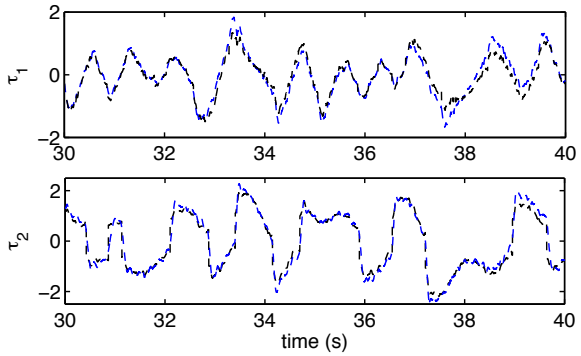


Fig. 4. Measured (dashed) and predicted (solid) torques for cross validation trajectory with non-SI units

III. LPV MODEL

A. Derivation of a Quasi-LPV Model

Consider an LPV model in the state-space form

$$\begin{aligned} \dot{x}(t) &= A_P(\boldsymbol{\theta}(t))x(t) + B_P(\boldsymbol{\theta}(t))u(t) \\ y(t) &= C_P(\boldsymbol{\theta}(t))x(t) + D_P(\boldsymbol{\theta}(t))u(t) \end{aligned}, \tag{7}$$

where $x \in \mathbb{R}^n$, $u \in \mathbb{R}^{n_u}$ and $y \in \mathbb{R}^{n_y}$. The mappings $A_P(.), B_P(.), C_P(.)$ and $D_P(.)$ are continuous functions of time-varying scheduling parameter vector $\boldsymbol{\theta}(t) \in \mathbb{R}^l$. In this problem, $n = 4$, $n_u = 2$, $n_y = 2$. This model can also be represented by a linear input-output map

$$P(\boldsymbol{\theta}) = \left[\begin{array}{c|c} A_P(\boldsymbol{\theta}) & B_P(\boldsymbol{\theta}) \\ \hline C_P(\boldsymbol{\theta}) & D_P(\boldsymbol{\theta}) \end{array}\right]. \tag{8}$$

The parameter vector $\boldsymbol{\theta}(t)$ depends on measurable signals $\boldsymbol{\rho}(t) \in \mathbb{R}^s$ referred to as scheduling signals, according to

$$\boldsymbol{\theta}(t) = f(\boldsymbol{\rho}(t)), \tag{9}$$

where $f : \mathbb{R}^s \to \mathbb{R}^l$ is a continuous mapping.

Consider the compact set $\mathcal{P}_{\boldsymbol{\theta}} \subset \mathbb{R}^l : \boldsymbol{\theta} \in \mathcal{P}_{\boldsymbol{\theta}}, \forall t > 0$. Here, it is assumed to be a polytope defined by the convex hull

$$\mathcal{P}_{\boldsymbol{\theta}} := \text{Co}\{\boldsymbol{\theta}_{v_1}, \boldsymbol{\theta}_{v_2}, \ldots, \boldsymbol{\theta}_{v_L}\}, \tag{10}$$

where $L = 2^l$ is the number of vertices.

The LPV system is called parameter-affine, if the state space model depends affinely on the parameters

$$P(\boldsymbol{\theta}) = \sum_{i=0}^{l} \theta_i P_i = P_0 + \theta_1 P_1 + \cdots + \theta_l P_l. \tag{11}$$

Since $\boldsymbol{\theta}$ can be expressed as a convex combination of L vertices θ_{v_i}, if (11) holds, it follows that the system can be represented by a linear combination of LTI models at the vertices; this is called a polytopic LPV system

$$P(\boldsymbol{\theta}) \in \text{Co}\{P(\theta_{v_1}), P(\theta_{v_2}), \ldots, P(\theta_{v_L})\} = \sum_{i=1}^{L} \alpha_i P(\theta_{v_i}), \tag{12}$$

where $\sum_{i=1}^{L} \alpha_i = 1$, and $\alpha_i \geq 0$ are the convex coordinates. To obtain the quasi-LPV model of the robot, $\boldsymbol{\rho}(t)$ is selected as the state vector of the system, which consists of the transformed joint positions and velocities

$$\boldsymbol{\rho}(t) = x(t) = \begin{bmatrix} \tilde{q}_1 & \tilde{q}_2 & \dot{\tilde{q}}_1 & \dot{\tilde{q}}_2 \end{bmatrix}^\top. \tag{13}$$

The state matrices of the quasi-LPV model of the robot are obtained from (4) and (5) as

$$\begin{aligned} A_P &= \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ \theta_1 & \theta_2 & \theta_3 & \theta_4 \\ \theta_5 & \theta_6 & \theta_7 & \theta_8 \end{bmatrix}, \\ B_P &= \begin{bmatrix} 0 & 0 \\ 0 & 0 \\ b_7\theta_9 & -b_3\theta_{10} - b_2\theta_9 \\ b_8\theta_9 - b_3\theta_{10} & b_3\theta_{10} + b_1\theta_9 \end{bmatrix}, \\ C_P &= \begin{bmatrix} I & O \end{bmatrix}, D_P = O, \end{aligned} \tag{14}$$

where I and O denote identity and zero matrices of appropriate dimensions. In this problem $l = 10$ and the affine scheduling parameters $\boldsymbol{\theta}$ are given in the Appendix B. The representation of the new joint position variables $\tilde{q}_i$ in (3) facilitates the affine generation in (14). Affine dependence of the state matrices in $\boldsymbol{\theta}$ is different from [15] to avoid increasing the number of scheduling parameters due to the Coulomb friction terms.

Since $L = 1024$, this quasi-LPV model would require a number of 2049 LMIs [7] to be simultaneously solved for a polytopic LPV controller synthesis, which is obviously a numerically challenging problem. Moreover, an implemented controller would have on-line dependence on 1024 vertex controllers. For practical reasons, the number of vertices should therefore be decreased.

B. Parameter Set Mapping

Parameter set mapping is a systematic procedure to find tighter regions in the space of the scheduling parameters. Moreover, approximations of LPV models can be obtained which neglect insignificant directions in the mapped parameter space without ad-hoc model simplifications and parameter freezing. Using parameter set mapping allows a trade-off between the number of parameters and model accuracy in a straightforward way. Altogether, it will lead to a less conservative controller synthesis [10], [15]. This method is used in this paper to reduce the computation cost and conservatism of the controller. The objective is to find a mapping $g : \mathbb{R}^s \to \mathbb{R}^m$ such that $m \leqslant l$, and

$$\boldsymbol{\phi}(t) = g(\boldsymbol{\rho}(t)), \tag{15}$$

yields a model

$$\begin{aligned} \dot{x}(t) &= \hat{A}_P(\boldsymbol{\phi}(t))x(t) + \hat{B}_P(\boldsymbol{\phi}(t))u(t) \\ y(t) &= \hat{C}_P(\boldsymbol{\phi}(t))x(t) + \hat{D}_P(\boldsymbol{\phi}(t))u(t) \end{aligned}, \tag{16}$$

that provides a satisfactory approximation of (7). Finding a suitable integer m is an important issue. The first step is to generate typical trajectories of the scheduling signals such that all expected operating regions of the plant are covered. A multi-sine trajectory covering the whole operation range of the robot is used for this purpose with $N = 70000$ data points. Then, corresponding scheduling parameters are computed to generate the data matrix

$$\Theta = \begin{bmatrix} \boldsymbol{\theta}(0) & \boldsymbol{\theta}(T) & \dots & \boldsymbol{\theta}((N-1)T) \end{bmatrix} \in \mathbb{R}^{l \times N}, \tag{17}$$

where $T = 0.001$. To put the same weight on each θ_i, all rows of the data matrix are normalized such that each has zero mean and a unit standard deviation

$$\Theta_i^n = \mathcal{N}_i(\Theta_i),\ \Theta_i = \mathcal{N}_i^{-1}(\Theta_i^n). \tag{18}$$

Now, PCA [20] is applied to the normalized data. Introduce the singular value decomposition of Θ^n

$$\Theta^n = [U_s\ U_n] \begin{bmatrix} \Sigma_s & 0 & 0 \\ 0 & \Sigma_n & 0 \end{bmatrix} \begin{bmatrix} V_s^\top \\ V_n^\top \end{bmatrix}, \tag{19}$$

where $\Sigma_s = \text{diag}(\sigma_1 \cdots \sigma_m)$, $\Sigma_n = \text{diag}(\sigma_{m+1} \cdots \sigma_l)$, $U_s \in \mathbb{R}^{l \times m}$, $V_s \in \mathbb{R}^{N \times m}$, $U_n \in \mathbb{R}^{l \times (l-m)}$ and $V_n \in \mathbb{R}^{N \times (N-m)}$, and assume that U_s, Σ_s and V_s correspond to the m significant singular values, such that

$$\hat{\Theta}^n = U_s \Sigma_s V_s^\top \approx \Theta^n, \tag{20}$$

is a reasonable approximation of the given data. The fraction of total variation v_m is calculated to evaluate the accuracy of the approximated model as

$$v_m = \frac{\sum_{i=1}^{m} \sigma_i^2}{\sum_{i=1}^{l} \sigma_i^2}, \tag{21}$$

where σ_i denote the singular values in (19). By choosing the number m of scheduling parameters, one can trade the accuracy of the model against complexity. The matrix U_s represents a basis of the significant column space of the data matrix Θ^n, and can be used to obtain a reduced mapping g in (15) by computing

$$\boldsymbol{\phi}(t) = g(\boldsymbol{\rho}(t)) = U_s^\top \mathcal{N}(f(\boldsymbol{\rho}(t))). \tag{22}$$

The approximated model in (16) is related to (7) by

$$P(\boldsymbol{\phi}) = \left[\begin{array}{c|c} \hat{A}_P(\boldsymbol{\phi}) & \hat{B}_P(\boldsymbol{\phi}) \\ \hline \hat{C}_P(\boldsymbol{\phi}) & \hat{D}_P(\boldsymbol{\phi}) \end{array}\right] = \left[\begin{array}{c|c} A_P(\hat{\boldsymbol{\theta}}) & B_P(\hat{\boldsymbol{\theta}}) \\ \hline C_P(\hat{\boldsymbol{\theta}}) & D_P(\hat{\boldsymbol{\theta}}) \end{array}\right], \tag{23}$$

$$\hat{\boldsymbol{\theta}}(t) = \mathcal{N}^{-1}(U_s \boldsymbol{\phi}(t)) = \mathcal{N}^{-1}(U_s U_s^\top \mathcal{N}(\boldsymbol{\theta}(t))), \tag{24}$$

where $\mathcal{N}^{-1}$ denotes the row-wise re-scaling. The approximated LPV model can be produced at any time by (24).

The above procedure is applied to the derived LPV model in (7). It turns out that about 57% of the information is contributed by the first principle component. Around 96% and 98% of the information is selected by choosing $m = 2$ and $m = 3$, respectively. In this paper $m = 2$ is chosen since it gives only 4 vertices and 9 LMIs to be solved for LPV controller synthesis. It is also obvious that the mapped parameter space with two dimensions has much less overbounding than the original one, leading to a less conservative controller. Some typical parameter trajectories for the accurate and approximated model are compared in Fig. 5, which illustrates a satisfactory accuracy for controller design purpose. Based on the derived quasi-LPV model, a polytopic LPV controller is designed.

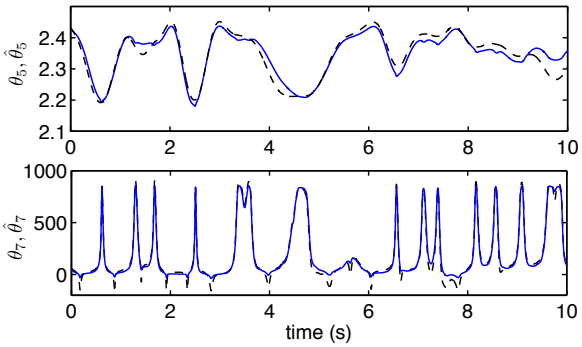


Fig. 5. Scheduling parameter θ_i (dashed) and approximation $\hat{\theta}_i$ (solid)

IV. CONTROLLER SYNTHESIS

Polytopic gain-scheduled controller design with a fixed Lyapunov function [7] has been proven to be an effective and practical tool for LPV synthesis due to the simplicity of the synthesis and implementation and low online computational effort. The reduced quasi-LPV model in the previous section is used for controller synthesis. The manipulator is controlled by an LPV controller that is scheduled by the reduced parameter vector. The design objective considered here is to stabilize the plant in the whole operating range with a high tracking capability, disturbance and measurement noise rejection and taking in consideration the actuator constraints. An $\mathcal{H}_\infty$ mixed-sensitivity loop-shaping approach is adopted to achieve the objectives. After discretizing the LPV model, a 12^{th} order discrete-time polytopic gain-scheduled $\mathcal{H}_\infty$ controller is synthesized similar to the design in [15]. The design procedure is not explained here again due to space limitation.

V. EXPERIMENTAL RESULTS

This section describes the experimental setup and the implementation of the designed LPV controller, and comparison of results with a similar LPV controller based on a linear friction model and two classical controllers. A low-complexity controller structure and reduced conservatism lead to an experimental implementation of the designed controller with a high performance. A low online computational load required for controller implementation allows to chose a high sampling frequency.

A. Experimental Setup

The CRS A465 industrial robotic manipulator has six rotational joints actuated by DC motors. The angular displacements of the motor shafts are measured by incremental encoders with a resolution of 1.5×10^{-5} rad. The robot has a repeatability of 0.05 mm with 2 Kg payload, and its harmonic drive transmission provides a smooth motion with zero backlash for all joints.

The system is supplied by the manufacturer with a C500C controller that controls the robot with a decentralized PD controller. It also connects both encoders and motor amplifiers to a PC-based open architecture control system via a Q8 input/output board of Quanser Consulting. Online data acquisition and control in MATLAB/SIMULINK is possible using a WinCon real-time Windows application [21]. The sampling frequency during the identification experiment and controller implementation is 1 kHz. Higher sampling rate is not possible due to hardware limitations.

B. Controller Implementation

The designed LPV controller which is referred to as LPV-CV is implemented on the two assigned joints of the robot, to track a specific trajectory which is different from identification trajectories. Joint velocities are obtained online by numerical differentiation of the joint positions after suitable low-pass filtering, since they are needed to compute the scheduling signals. For comparison, a similar LPV controller based on a linear viscous friction model [15] referred to as LPV-V, an inverse dynamics feedforward controller based on an accurate identified model in [19] and a decentralized PD controller proposed by the manufacturer are also implemented.

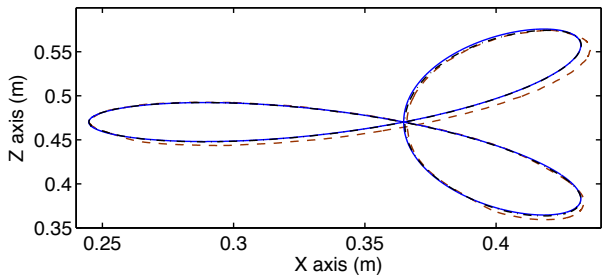


Fig. 6. Cartesian command trajectory (solid), LPV-CV (dashed-dotted) and PD (dashed) controller outputs

The trajectory tracking results in Cartesian space are plotted in Fig. 6 only for LPV-CV and PD controllers for a convenient comparison (Other plots are indistinguishable in print). The joint space trajectories together with the reference inputs for LPV-CV and PD controllers are plotted in Fig. 7, where the figures are zoomed in to make the deviations more visible. The root mean square (RMS) of tracking error for the whole trajectory by all controllers are given in Table II.

The experimental results in both Cartesian and joint spaces illustrate the high accuracy of the designed LPV-CV controller. It has a better tracking performance compared with

TABLE II

RMS OF TRACKING ERROR FOR ALL IMPLEMENTED CONTROLLERS IN CARTESIAN SPACE (METER) AND JOINT SPACE (DEGREE)

Controllers	Cartesian Space	Joint Space
LPV-CV	$3.42\ 10^{-4}$	0.0481
LPV-V	$3.66\ 10^{-4}$	0.0516
Inverse Dynamics	$3.71\ 10^{-4}$	0.0522
PD	$3.0\ 10^{-3}$	0.4084

the LPV-V controller. Moreover, the LPV-CV and LPV-V controller obviously outperform the PD controller by a factor of 8 in terms of tracking performance and achieve a slightly better accuracy than a model-based inverse dynamics controller. Since the inverse dynamics control law includes the whole dynamic terms of the robot, it is more complex especially for higher DOFs.

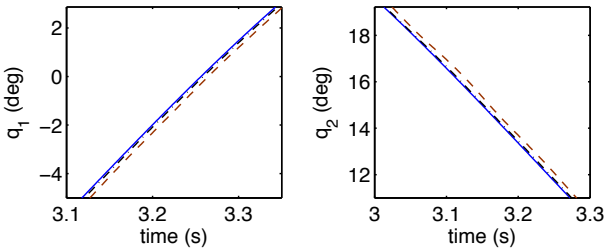


Fig. 7. Joint command trajectory (solid), LPV-CV (dashed-dotted) and PD (dashed) controller outputs for both joints (zoomed-in)

The tracking experiment is done for 13 seconds, including two periods of the Cartesian trajectory tracking plus initial and final smoothing trajectories. The control signal computed by the LPV controller during the tracking experiment is far below the saturation limits but not shown here due to space limitation.

VI. CONCLUSIONS

This paper presents a realistic LPV modelling and control of a 2-DOF robotic manipulator. All rigid-body dynamic and important friction terms of the manipulator have been taken into account in the modelling without any ad-hoc model simplification. The adopted friction model includes the Coulomb and viscous terms , where the hard nonlinearity in the Coulomb term has been approximated by a hyperbolic tangent function. Since the quasi-LPV system has a large number of scheduling parameters and a high overbounding, the systematic procedure of PCA-based parameter set mapping has been used to find tighter scheduling parameter set with fewer dimensions, which results in a less conservative controller with lower online computation load. The parameter set dimension has been reduced from 10 to 2 by this mapping, reducing the number of LMIs to be solved for controller synthesis from 2049 to 9.

A polytopic LPV gain-scheduled controller has been synthesized based on the approximated LPV model and is

implemented successfully on an industrial robot. It is illustrated that the designed LPV controller has a better tracking performance compared with a similar LPV controller based on a linear viscous friction model. Moreover, it obviously outperform a PD controller and achieves a slightly better accuracy than a model-based inverse dynamics controller. The advantage of the LPV controller over the inverse dynamics one is that it is less complex and various design objectives can be adopted according to the performance and robustness requirements of the application.

The computed control signal by the LPV controller is far below the actuator saturation limits, and the high measurement noise of encoders are rejected since both facts are considered in the mixed-sensitivity loop-shaping design.

If the setup is equipped with high resolution position sensors, dynamic friction models including presliding regime can also be used in LPV modelling, following the same procedure proposed in this paper.

APPENDIX

A. Grouped Dynamic and Kinematic Parameters

$$
\begin{aligned}
&b_1 = I_{zz_1} + m_2 a_1^2, \quad b_2 = I_{zz_2}, \quad b_3 = M_{y_2} a_1,\\
&b_4 = g M y_2, \quad b_5 = g(m_2 a_1 + M x_1), \quad b_6 = f_{v_1},\\
&b_7 = I_{m_2} + I_{zz_2}, \quad b_8 = I_{m_2}, \quad b_9 = f_{v_2},\\
&b_{10} = f_{c_1}, \quad b_{11} = f_{c_2},
\end{aligned}
$$

where I_{zz_n} is the moment of inertia of *n*th link, m_n and a_n are the mass and length of the links, M_{x_n} and M_{y_n} are the first moment of the *n*th link (product of mass and coordinates of gravity center), and I_{m_n} is the motor moment of inertia. The axes are named according to the modified DH convention.

B. Scheduling Parameters

$$
\begin{aligned}
\theta_1 &= b_7 b_5 \mathrm{sinc}(\tilde{q}_1)\frac{1}{\nu},\\
\theta_2 &= \mathrm{sinc}(\tilde{q}_2)(b_7 b_4 - b_2 b_4 - b_3 b_4 \cos(\tilde{q}_1 - \tilde{q}_2))\frac{1}{\nu},\\
\theta_3 &= (-b_7 b_6 - b_2 b_3 \sin(\tilde{q}_1 - \tilde{q}_2)\dot{\tilde{q}}_1 - b_2 b_9 - b_3 b_9 \cos(\tilde{q}_1 - \tilde{q}_2)\\
&\quad - b_3^2 \cos(\tilde{q}_1 - \tilde{q}_2)\sin(\tilde{q}_1 - \tilde{q}_2)\dot{\tilde{q}}_1 - b_7 b_{10}\mathrm{tanhc}(\dot{\tilde{q}}_1)\\
&\quad - b_3 b_{11}\cos(\tilde{q}_1 - \tilde{q}_2)\mathrm{tanhc}(\dot{\tilde{q}}_2 - \dot{\tilde{q}}_1) - b_2 b_{11}\mathrm{tanhc}(\dot{\tilde{q}}_2 - \dot{\tilde{q}}_1)\\
&\quad + b_7 b_3 \sin(\tilde{q}_1 - \tilde{q}_2)\dot{\tilde{q}}_1)\frac{1}{\nu},\\
\theta_4 &= (b_2 b_9 - b_7 b_3 \sin(\tilde{q}_1 - \tilde{q}_2)\dot{\tilde{q}}_2 + b_3 b_9 \cos(\tilde{q}_1 - \tilde{q}_2)\\
&\quad + b_{11}(b_2 + b_3 \cos(\tilde{q}_1 - \tilde{q}_2))\mathrm{tanhc}(\dot{\tilde{q}}_2 - \dot{\tilde{q}}_1))\frac{1}{\nu},\\
\theta_5 &= (-b_3 b_5 \cos(\tilde{q}_1 - \tilde{q}_2)\mathrm{sinc}(\tilde{q}_1) + b_8 b_5 \mathrm{sinc}(\tilde{q}_1))\frac{1}{\nu},\\
\theta_6 &= ((b_8 + b_1)b_4 \mathrm{sinc}(\tilde{q}_2))\frac{1}{\nu},\\
\theta_7 &= (b_1 b_9 + b_3 b_9 \cos(\tilde{q}_1 - \tilde{q}_2) - b_8 b_6 + b_3 b_6 \cos(\tilde{q}_1 - \tilde{q}_2)\\
&\quad + b_8 b_3 \sin(\tilde{q}_1 - \tilde{q}_2)\dot{\tilde{q}}_1 + b_1 b_3 \sin(\tilde{q}_1 - \tilde{q}_2)\dot{\tilde{q}}_1\\
&\quad + b_3 b_{11}\cos(\tilde{q}_1 - \tilde{q}_2)\mathrm{tanhc}(\dot{\tilde{q}}_2 - \dot{\tilde{q}}_1) + b_1 b_{11}\mathrm{tanhc}(\dot{\tilde{q}}_2 - \dot{\tilde{q}}_1)\\
&\quad + b_3 b_{10}\cos(\tilde{q}_1 - \tilde{q}_2)\mathrm{tanhc}(\dot{\tilde{q}}_1) - b_8 b_{10}\mathrm{tanhc}(\dot{\tilde{q}}_1))\frac{1}{\nu},\\
\theta_8 &= (-b_1 b_9 + b_3^2 \cos(\tilde{q}_1 - \tilde{q}_2)\sin(\tilde{q}_1 - \tilde{q}_2)\dot{\tilde{q}}_2\\
&\quad - b_3 b_9 \cos(\tilde{q}_1 - \tilde{q}_2) - b_8 b_3 \sin(\tilde{q}_1 - \tilde{q}_2)\dot{\tilde{q}}_2\\
&\quad - (b_1 - b_3 \cos(\tilde{q}_1 - \tilde{q}_2))b_{11}\mathrm{tanhc}(\dot{\tilde{q}}_2 - \dot{\tilde{q}}_1))\frac{1}{\nu},\\
\theta_9 &= \frac{1}{\nu},\\
\theta_{10} &= \cos(\tilde{q}_1 - \tilde{q}_2)\frac{1}{\nu},\\
\nu &= b_7 b_1 + b_2 b_8 +\\
&\quad + \cos(\tilde{q}_2 - \tilde{q}_1)(b_7 b_3 - b_2 b_3 + b_3 b_8 - \cos(\tilde{q}_2 - \tilde{q}_1)b_3^2),
\end{aligned}
$$

where $\mathrm{sinc}(x) = \frac{\sin(x)}{x}$ and $\mathrm{tanhc}(x) = \frac{\tanh(\beta x)}{x}$.

REFERENCES

[1] B. Armstrong-Helouvry, *Control of machines with friction*. Kluwer Academic Publishers, Boston, USA, 1991.

[2] H. Olsson, K. J. Astrom, C. Canudas de Wit, M. Gafvert, and P. Lischinsky, "Friction Models and Friction Compensation," *European Journal of Control*, vol. 4, no. 3, pp. 176–195, 1998.

[3] A. Harnoy, B. Frieland, and S. Cohn, "Modeling and Measuring Friction Effects," *IEEE Control Systems Magazine*, vol. 28, no. 6, pp. 82–91, 2008.

[4] C. Canudas de Wit, H. Olsson, K. Aström, and P. Lischinsky, "A new model for control of systems with friction," *IEEE Transactions on Automatic Control*, vol. 40, no. 3, pp. 419–425, 1995.

[5] J. Swevers, F. Al-Bender, C. G. Ganseman, and T. Prajogo, "An integrated friction model structure with improved presliding behavior for accurate friction compensation," *IEEE Trans on Automatic Control*, vol. 45, no. 4, pp. 675–686, 2000.

[6] Y. Kunii, B. Solvang, G. Sziebig, and P. Korondi, "Tensor Product Transformation Based Friction Model," in *Proc. of the 11th International Conference on Intelligent Engineering Systems*, Hungary, 2007, pp. 259–264.

[7] P. Apkarian, P. Gahinet, and G. Becker, "Self-scheduled H_∞ control of linear parameter-varying systems: a design example," *Automatica*, vol. 31, no. 9, pp. 1251–1261, 1995.

[8] P. Apkarian and P. Gahinet, "A convex characterization of gain scheduled $\mathcal{H}_\infty$ controllers," *IEEE Transactions on Automatic Control*, vol. 40, pp. 853–864, 1995.

[9] F. Wu and K. Dong, "Gain-scheduling control of LFT systems using parameter-dependent Lyapunov funtions," *Automatica*, vol. 42, no. 1, pp. 39–50, 2006.

[10] A. Kwiatkowski and H. Werner, "PCA-based parameter set mappings for LPV models with fewer parameters and less overbounding," *IEEE Trans. on Control Systems Technology*, vol. 16, no. 4, pp. 781–788, 2008.

[11] C. Vivas and F. R. Rubio, "Gain-scheduling Control of Systems with Dynamic Friction," in *Proc. 41st IEEE Conference on Decision and Control*, 2002, pp. 89–94.

[12] Z. Yu, H. Chen, and P. Woo, "Gain scheduled LPV H_∞ control based on LMI approach for a robotic manipulator," *Journal of Robotic Systems*, vol. 19, no. 12, pp. 585–593, 2002.

[13] A. Kwiatkowski and H. Werner, "LPV control of a 2-DOF robot using parameter reduction," in *Proc. 44th IEEE Conference on Decision and Control and European Control Conference*, Seville, Spain, 2005, pp. 3369–3374.

[14] H. S. Ali, L. Boutat-Baddas, Y. Becis-Aubry, and M. Darouach, "H_∞ control of a scara robot using polytopic lpv approach," in *14th Mediterranean Conference on Control and Automation*, Italy, 2006.

[15] S. M. Hashemi, H. Abbas, and H. Werner, "LPV modelling and control of a 2-DOF robotic manipulator using PCA-based parameter set mapping," in *Proc. 48th IEEE Conference on Decision and Control*, Shanghai, China, 2009, pp. 7418–7423.

[16] H. Abbas, S. M. Hashemi, and H. Werner, "Decentralized LPV gain-scheduled PD controller of a robotic manipulator," in *Proc. ASME Dynamic Systems and Control Conference*, California, USA, 2009.

[17] M. Daemi and B. Heimann, "Identification and compensation of gear friction for modeling of robots," in *Proc. of the 11th CISM-IFToMM Symp. Theory and Practice Robots and Manipulators*, Udine, Italy, 1996, pp. 89–96.

[18] J. SWEVERS, W. VERDONCK, and J. D. SCHUTTER, "Dynamic model identification for industrial robots," *IEEE Control Systems Magazine*, pp. 58–71, 2007.

[19] S. M. Hashemi and H. Werner, "Parameter identification of a robot arm using separable least squares technique," in *Proc. European Control Conference*, Budapest, Hungary, 2009, pp. 2199–2204.

[20] J. Jackson, *A user's guide to principle components*, ser. Wiley series in probabilitiy and mathematical statistics. John Wiley & Sons. inc, 1991.

[21] Quanser Consulting, "Wincon 4.1 user's guide," 2003.

10 An Observer Based Scheme for Adapt to Blade Aerodynamic Parameters for Power Control of Wind Turbines

An Observer Based Scheme for Adapt to Blade Aerodynamic Parameters for Power Control of Wind Turbines

Peter Fogh Odgaard
kk-electronic a/s
Viby J, Denmark
peodg@kk-electronic.com

Rasmus Nielsen
kk-electronic a/s
Ikast, Denmark
ranie@kk-electronic.com

Chris Damgaard
kk-electronic a/s
Ikast, Denmark
chdam@kk-electronic.com

***Abstract*— This paper presents an observer based scheme for adapting the power control of wind turbines to the actual power coefficients of the blades mounted on the wind turbine. Normally it is assumed that the power coefficients for one turbine in a production series are valid for all the turbines in that production series. An unknown input observer is used to estimate the actual table of power coefficients depending on blade pitch angle and tip speed ratio. If the actual table is much different from the initial assumed table, the actual table is found by iterations. A simulation is used to illustrate the schemes potential to estimate the power coefficients and to see the gained potential of the use of this scheme compared with a non-corrected situation.**

Keywords- Wind turbines; unknown input observers; adaptive systems; aerodynamics; power control

I. INTRODUCTION

In the recent years the installed energy generation capacity of wind turbines has dramatically increased. In the same period the sizes of the individual wind turbines has increased from turbines in the kilo watt sizes to multiple mega watts. Consequently optimization of the power generated by the turbine is as well of increasingly interest.

A wind turbine consists of a tower on which a nacelle is mounted; in the nacelle the generator is placed. The generator is driven by a main shaft at which the turbine's blades are fixed. In the example used in this paper a turbine with 3 pitch-able blades are used, meaning that their angle towards the wind can be controlled. A gearbox, enabling the possibility of different rotational speeds, divides the main shaft. The increased size of the wind turbines have also increased the interest of turbines with variable speed and active pitch of the blades, which are used to keep the turbine at rated power when the rated wind speed is exceeded. Until the rated power is achieved the power optimum is obtained by requiring the optimal reference torque at the generator, see [1]

Some research has been conducted on the subject of optimization of the generated power of the wind turbine. A couple of examples can be found in [1] and [2].

The optimal torque and pitch references are obtained by a mapping between power generation ratio, pitch angle and the tip speed ratio which is the ratio between wind speed and the speed of the blade tip, (the rotation speed of the rotor can be controlled by the torque reference). This mapping is denoted the C_p -surface and could be obtained by: finite element simulations, wind tunnel experiments etc., in practice it is often represented by a table, which in the following is denoted as the C_p-table.

A problem in achieving optimal power and speed control of a wind turbine is that the C_p-surface is not well known. Initially these values are most often not actually measured but computed, and if measured, only a few blades in an entire production series and measurements are rarely performed on the actual turbine. Some work has been published regarding adapting the power controller to the specific C_p-surface. In [1] a scheme is proposed which uses a Newton like scheme to find the power optimum online. A non-linear controller is proposed in [2], which assumes that the wind speed is well known. [3] presents an adaptive scheme which uses a least square method to online identify the system parameters; the controller is designed using a minimum variance controller, which is impractical. [4] presents a fuzzy control to adapt the power controller.

In [5] another part of the solution dealing with this problem is presented. It presents a method to estimate the C_p-value and the wind speed online based on standard measurements from the wind turbine. The estimated C_p-values can subsequently after approximately 1 month be used to update of the C_p-surface. Based on the updated C_p -surface a new power optimum can be found. This means that the power control is adapted using an adaptive C_p-surface. A large advantage of this scheme is that the existing control structure is not influenced by this scheme, it do only provide updated power references when present. An unknown input observer is used to estimate these variables, for details on the general scheme see [6], in this

context it is changed to estimate the unknown inputs instead of being robust towards them, another example can be found in [7].

Based on these online estimations of the C_p-surface a simple scheme was proposed to adapt the optimal reference point to the power controller of wind turbine, in respect with changes in the C_p-surface due to change with time e.g. due to debris build-up on the blades of the turbine, see [8]. In this paper the proposed scheme will both be extended and also bee changed to handle the estimation of the real C_p-surface and its optimum point. This estimation is designed to be performed during the commission of the wind turbines.

The model of the wind turbine is subsequently introduced, which leads to an estimator design using the optimal unknown input observer scheme. Based on the estimates of the C_p-values a C_p-table can be computed, which again can be used to find the optimum of the C_p-table. This optimum is used as reference for the power controller of the wind turbine.

In Section II a wind turbine model is described. The proposed scheme is presented in Section III and Section IV. In Section V the end the simulation is used to show the potential gain of using this proposed adaptive scheme to adapt the actual optimum of the C_p-table, by comparing a wind turbine with and without an adapted optimum reference value. A conclusion is drawn in Section VI.

II. Wind Turbine Model

In a typical variable speed wind turbine the generated power is controlled by two modes power and speed control. In power control mode the generator torque is controlled to maximize the generated power, by obtaining a tip speed ratio which results in the optimum on the C_p-surface. The produced power of these two control modes is mapped as a function of the wind speed in Fig.1. In speed control mode the blades are pitched such that the rated power is obtained and the generator torque chosen such that the rotational speed of the rotor is following the nominal value. In order to obtain the optimum power in power control mode the C_p -surface is highly important, e.g. the turbine will not be controlled optimally if the optimum on C_p-surface is moved from the assumed.

Inspecting the problem deeper, the torque balance model of the wind turbine is considered.

$$\dot{\omega}(t) = \frac{1}{J}\Big(\tau_{\text{aero}}(t) - \tau_{\text{ref}}(t)\Big), \tag{1}$$

and

$$\tau_{\text{aero}}(t) = \frac{\rho A C_p\big(\theta(t),\lambda(t)\big)v(t)^3}{2\omega(t)}, \tag{2}$$

where $\omega(t)$ is the rotational velocity of the rotor, $\tau_{\text{ref}}(t)$ is the reference torque to the generator transferred to the low speed side, and due to the fast power electronics in the converter and the generator it is assumed that this reference is followed, $\theta(t)$ is the pitch angle, $\lambda(t)$ is the tip speed ratio, $v(t)$ is the wind speed, J is the moment of inertia of blades shaft etc, ρ is the density of the air, A is the area covered by the blades in the rotation. $\omega(t)$ and $\tau_{\text{ref}}(t)$ are measurable. The wind denoted $v(t)$ is measured as well, but is very non-reliable, since it should be the average over the entire swept area, and not a point measurement, in addition the wind is measured with an anemometer mounted on the wind turbine nacelle behind the blades. The aerodynamics of the blades will consequently influence the wind speed measurement by the anemometer.

If this model is linearized a small signal model can be obtained, the variations between actual power coefficients and wind speed compared to measured and computed once are denoted as $\Delta C_p(t)$ and $\Delta v(t)$.

$$\begin{aligned}\Delta\tau_{\text{aero}}(t) = &-\frac{\rho\cdot A\cdot C_{p,0}\cdot v_0^3}{2\cdot\omega_0^2}\Delta\omega(t) + \frac{\rho\cdot A\cdot v_0^3}{2\cdot\omega_0}\Delta C_p(t)\\ &+\frac{3\cdot\rho\cdot C_{p,0}\cdot v_0^2}{2\cdot\omega_0}\Delta v(t).\end{aligned} \tag{3}$$

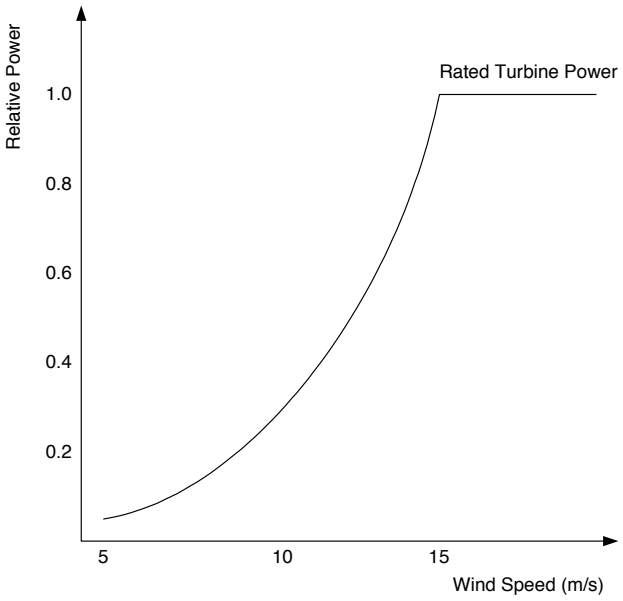


Fig. 1 Illustration of the produced power until the rated power is reached at a wind speed of 15 m/s, the power is optimized and above this speed it is limited by blade pitching.

In this context the power coefficient and the wind speed is assumed to be unknown variables, which varies around a well-known working point, as modelled by the small signal values of these. Changes in the C_p- table would be very slowly, meaning that the frequency content of $\Delta C_p(t)$ is in the region close to 0 rad/s. Compared to this the changes in the wind speed will be placed in a region with much higher frequency content, meaning that frequency separation can be assumed.

The anemometer, introduces a risk of a DC-error on the measurement. This is much more critical than high uncertainties at higher frequencies since these can be decoupled by the frequency information. If a DC calibration of the anemometer is not performed, an offset will be introduced on

the C_p-table estimates. It will be a constant offset on the entire C_p-table, and consequently not resulting in a false optimal C_p-value, but the absolute value cannot be determined.

III. Observer Design

In order to design the observer model of the wind turbine, the model should be extended similar to the method used in [7].

The existence of the unknown inputs in the system, points at the usage of a specific scheme to estimate the C_p- values, this is the optimal unknown input observer. In order to use this unknown input observer $\Delta C_p(t)$ is used as a state; it can be modeled as low pass filter of first order. However, used in the commissioning process of the wind turbine this means that the C_p-tables should be recomputed a number of times in order to iterate to the correct table. In order to use this approach the wind speed needs to be estimated as well, and it can be assumed to change much faster than the C_p-table, meaning it can be represented by a band pass filter, the difference between the measured and actual wind speed was previously defined as $\Delta v(t)$. In principle these filters represents the uncertainties in the C_p-table and the wind speed. In practice the difference in the relevant frequencies is relatively large. In terms of time durations of the relevant content are days for the C_p-table and in the milliseconds range for the wind speed variations.

The linear wind turbine model extended with the internal models of $\Delta C_p(t)$ and $\Delta v(t)$ can be seen in

$$\begin{aligned}\dot{\mathbf{x}}_c(t) =&\mathbf{A}_c\mathbf{x}_c(t) + \mathbf{B}_c\mathbf{u}_c(t) + \mathbf{E}_c\mathbf{d}_d(t),\\ \mathbf{y}_c(t) =&\mathbf{C}_c\mathbf{x}_c(t),\end{aligned} \tag{4}$$

where

$$\mathbf{x}_c(t) = \begin{bmatrix}\omega(t)\\ \mathbf{x}_{C_p}(t)\\ \mathbf{x}_v(t)\end{bmatrix},\ \mathbf{y}_c(t) = \begin{bmatrix}\tau_{\text{aero}}(t)\\ \omega(t)\\ v(t)\end{bmatrix}, \mathbf{u}_c(t) = \tau_{\text{ref}}(t),$$

$\mathbf{x}_{C_p}(t)$ is a state vector representing the internal model of $\Delta C_p(t)$, $\mathbf{x}_v(t)$ is the state vector representing the internal model of $\Delta v(t)$, $d_d(t)$ is a signal representing the unknown input. The internal model of the $\Delta C_p(t)$ is in state space form $(\mathbf{A}_{C_p},\ \mathbf{B}_{C_p},\ \mathbf{C}_{C_p},\ \mathbf{D}_{C_p})$, and the internal model of $\Delta v(t)$ is in state space form $(\mathbf{A}_v,\ \mathbf{B}_v,\ \mathbf{C}_v,\ \mathbf{D}_v)$, the merged system matrices are defined by:

$$\mathbf{A}_c = \begin{bmatrix}-\frac{\rho A C_{p,0} v_0^3}{2\omega_0^2 J} & \frac{\rho A v_0^3}{2\omega_0 J}\mathbf{C}_{C_p} & \frac{3\rho A C_{p,0} v_0^2}{2\omega_0 J}\mathbf{C}_v\\ 0 & \mathbf{A}_{C_p} & 0\\ 0 & 0 & \mathbf{A}_v\end{bmatrix}, \tag{5}$$

$$\mathbf{B}_c = \begin{bmatrix}\frac{-1}{J}\\ 0\\ 0\end{bmatrix}, \tag{6}$$

$$\mathbf{C}_c = \begin{bmatrix}1 & 0 & 0\\ 0 & v_0^3\mathbf{C}_{C_p} & 3C_{p,0}v_0^2\mathbf{C}_v\\ 0 & 0 & \mathbf{C}_v\end{bmatrix}, \tag{7}$$

$$\mathbf{E}_c = \begin{bmatrix}10^{-3}\\ \mathbf{B}_{C_p}\\ \mathbf{B}_v\end{bmatrix}, \tag{8}$$

The first element in $\mathbf{E}_c$ is non-zero even though that $\omega(t)$ is not assumed to be driven by the uncertain input. The reason is that this small but non-zero elements in the matrix introduce some robustness towards model uncertainties, which could be due to linearization of the nonlinear model before it is used to design the observer.

The system is subsequently descretized with a sample frequency at 10 Hz. Resulting in a state space system defined as in (9), where stochastic disturbances and measurement noises are added.

$$\begin{aligned}\mathbf{x}_c[n+1] =&\mathbf{A}_d\mathbf{x}_c[n] + \mathbf{B}_d\mathbf{u}_c[n] + \mathbf{E}_d d_d[n] + \zeta[n],\\ \mathbf{y}_c[n] =&\mathbf{C}_d\mathbf{x}_c[n] + \eta[n],\end{aligned} \tag{9}$$

in addition to the filters defining the internal models, the covariance matrices Q and R introduces design flexibility into the system as a couple of design parameters. The filters representing $\Delta C_p(t)$ and $\Delta v(t)$ have to be designed, the point is that $\mathbf{A}_{C_p}$, $\mathbf{B}_{C_p}$, and $\mathbf{C}_{C_p}$ are found such that it is a low pass filter with a time constant of days, and $\mathbf{A}_v$, $\mathbf{B}_v$ and $\mathbf{C}_v$ such that they form a high pass filter/ band pass filter such that its pass region is placed in the much higher region in terms of frequency content with a time constant in the milliseconds.

Due to the non-linearity of the model, especially the cubic dependency on the wind speed, a number of operating points are used in practice such that an observer is designed for each, and bump-less transfer is used to switch between them. However, for simplicity this is left out in this paper, and only a single observer is consequently designed to one point of operation. However, the introduction of these multiple observers would increase the performance of the estimated; but, as one can see in the simulation section, the scheme with only one observer performs pretty well.

A. Unknown Input Observer

The unknown input observer is given by a state space representation seen in (10).

$$\begin{aligned}\mathbf{z}[n+1] =&\mathbf{F}_{n+1}\mathbf{z}[n] + \mathbf{T}_{n+1}\mathbf{B}\mathbf{u}[n] + \mathbf{K}_{n+1}\mathbf{y}[n],\\ \hat{\mathbf{x}}[n+1] =&\mathbf{z}[n+1] + \mathbf{H}_{n+1}\mathbf{y}[n+1],\end{aligned} \tag{10}$$

In where the matrices are found using an algorithm described in [6].

B. Online Power Coefficients Table Computation

The estimated C_p-values are stored in an array depending on the pitch angle and tip speed ratio of the blades. The values are stored as the sum of the $\Delta C_p(t)$ and the assumed C_p-values based on the aerodynamic model. Each measurement is stored in the element with the closest geometrical origin; which is illustrated in Fig. 2, where the most recent measurement is marked with a black dot, and the nearest entry is (k-1,m).

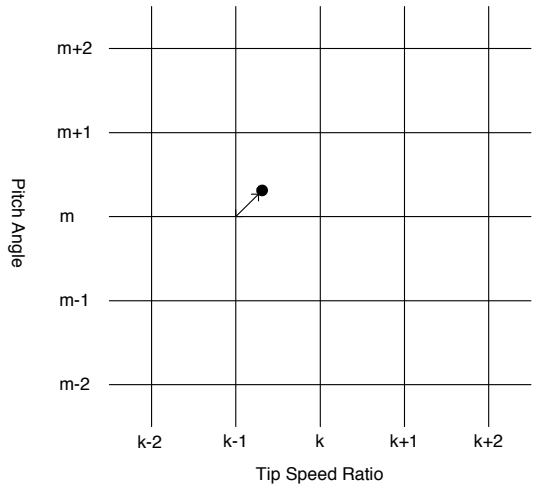


Fig. 2 Illustration of the location of the closets table entry. The measurement is marked with the black dot, and the nearest table entry (k-1,m).

In order to limit the needed memory for storing all the C_p-values, the mean value is computed iteratively, meaning that the computed mean of C_p-value and a counter of the C_p-values contained in the mean are stored.

The mean value of C_p-table at the point χ in the table is denoted as $\bar{C}_{p,\chi}$and defined as

$$\bar{C}_{p,\chi} = \frac{j_\chi \cdot \bar{C}_{p,\chi} + C_{p,\chi}}{j_\chi + 1}, \qquad (11)$$

Where j_χ is the counter of elements used to compute the old mean value. The computation of the C_p -table can be subsequently be defined as the following algorithm.

Find the element χ of the present C_p-value.

Compute a new mean for this element using (11).

Update the counter as

$j_\chi = j_\chi + 1$.

After the C_p-table has been computed for a sufficient time period, denote this time period as T , the table will be recomputed using the same measurement data as used to the present table, the only different is that the most recent computed C_p-table is used to compute the expected C_p- values used in the estimation. Continue this iterative process until the difference between the recent and previous C_p-table is small enough. The time period can be decreased if λ and β contains increased actuations.

Define the table error as

$$C_{p,\text{error}} = \Sigma_{\text{for}\cdots\chi\in\{\theta,\lambda\}} \left(\bar{C}_{p,\chi}[n] - \bar{C}_{p,\chi}[n-1]\right)^2, \qquad (12)$$

It is now possible to summarize the algorithm to compute.

Set n=1.

Compute first C_p-table using the previous algorithm, define this table as $\mathbf{C}_p[x],\ x = 1$.

Set $n = n + 1$.

Replace the initial $\mathbf{C}_p$with $\mathbf{C}_p[n+1]$, and compute $\mathbf{C}_p[n]$ as in 2).

Compute the table error as in (12). If $C_{p,\text{error}} < \nu$, stop the algorithm, else jump to 3).

The value of T should be found by experiments, it should at least represent so long time that the relevant parts of the table is well covered with a large enough numbers of instances. The table can covered faster if additional actuation signals are added to λ and β. An initial guess on the value of T is 1 month.

IV. Table Based Optimization

The optimization can be performed quite simple. Just find the point in the table where the maximal C_p-value is found. Since this algorithm is only performed once per time period computational efficiency is not very important. However, due the implementation in the wind turbine it is still relevant to limit the computational burden of the algorithm. In this first version of the algorithm a simple max search is used.

V. Simulations

In this simulation the same set of model parameter is used as in [5]. In which the observer used to estimate the C_p-values is designed.

The wind turbine is modelled by the nonlinear model in (1)-(2), and the following model parameters are used:

$J = 7.8 \cdot 10^6 \text{kg} \cdot \text{m}^2$, $R = 26\text{m}$, $\rho = 1.225 \frac{\text{kg}}{\text{m}^3}$

The sample frequency of the system and wind data is 10 Hz.

For the linear model the points of operation are chosen such that the entire range of the wind speed in the data set is covered as well as possible. The values are found to be $C_{p,0} = 0.46$, $v_0 = 15\frac{\text{m}}{\text{s}}$, $\omega_0 = 1\frac{\text{rad}}{\text{s}}$ and $\tau_{\text{ref},0} = 1.35 \cdot 10^6 \text{Nm}$.

The linear observer model matrices of the two internal models are found iteratively to:

$(A_{C_p} = -10^{-7},\ B_{C_p} = 1,\ C_{C_p} = 10^{-7},\ D_{C_p} = 0)$ and

$\left(\mathbf{A}_v = \begin{bmatrix} -1.01 & -0.10 \\ 0.10 & 0 \end{bmatrix},\ \mathbf{B}_v = \begin{bmatrix} 1 \\ 0 \end{bmatrix}\right)$,

$\left(\mathbf{C}_v = \begin{bmatrix} 1.00 & 0.001 \end{bmatrix},\ \mathbf{D}_v = 0\right)$.

The two cross correlation matrices are found iteratively as:

$$\mathbf{Q} = \mathbf{I}_{4\times 4},\ \mathbf{R} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 10^5 \end{bmatrix}.$$

In [5] this observer is used to estimate the C_p-values, the same simulations is performed with a change in the initial C_p-table to test how fast the observer will converge to the actual C_p-values, this shown in Fig. 3.

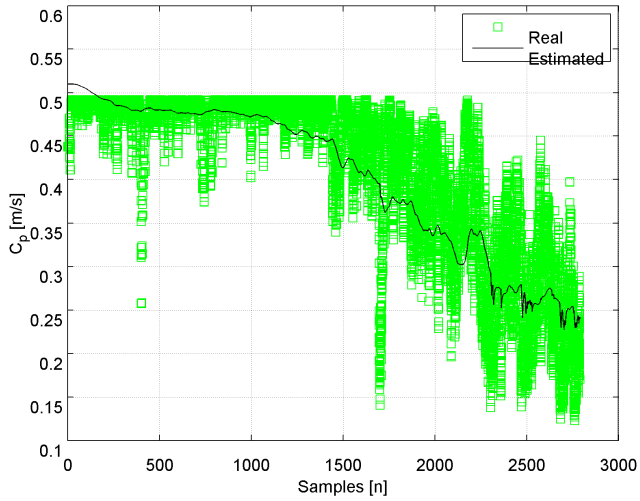


Fig. 3 Simulated C_p-values compared with the estimated one, simulated with an error in the initial C_p-value.

This simulation shows that the used observer estimates the C_p-values quite well even though the initial value of C_p-table is wrong.

Subsequently it is simulated how much can be gained by using the proposed adaptive scheme in a case where the optimal tip speed ratio λ is assumed to be 10% wrong. Since the pitch angle kept constant under power control of the wind turbine, C_p-table is in this simulation only depending on λ. These two C_p-curves can be seen in Fig. 4. From this figure it can be seen that the actual optimal λ-value is approximately equal 8 and that the assumed λ is approximately equal 7.

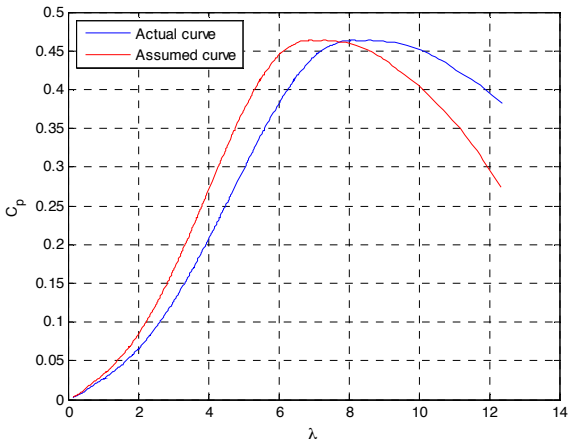


Fig. 4 Comparison of the actual C_p-curve only depending on λ and the assumed C_p-curve.

In order to simulate the non-adapted C_p-curve and the adapted C_p-curve measured wind data is used, see Fig. 5.

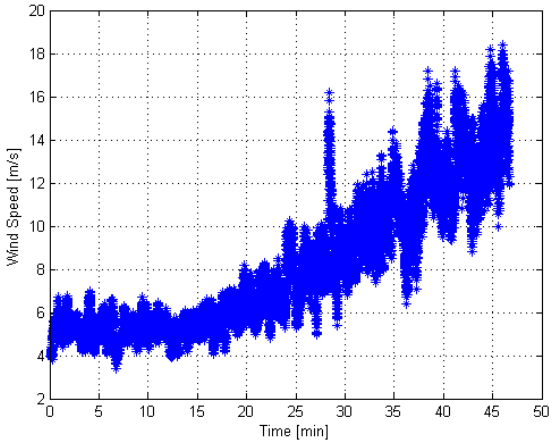


Fig. 5 The wind speed profile used in the simulation, which includes 47 minuets ranging from 5 m/s to 18 m/s.

The wind turbine is now simulated on this wind data, sampled at 10Hz, for both the adapted and non-adapted situation. A standard proportional controller is used as presented in [1], in which

$$\tau_{\text{ref}} = K\omega(t)^2, \tag{13}$$

where

$$K = \frac{1}{2}\rho A r^3 \frac{C_{p_{\max}}}{\lambda_*^3}, \tag{14}$$

in which $C_{p_{\max}}$ is the maximal power coefficient and which is equal 0.45, λ_* is the tip speed ratio relating to the maximal power coefficient, which in this case is 8 for the adapted case and 7 for the non-adapted. The produced powers for these simulations are both shown in Fig. 6. In the figure the powers are plotted relative to the rated power of the wind turbine.

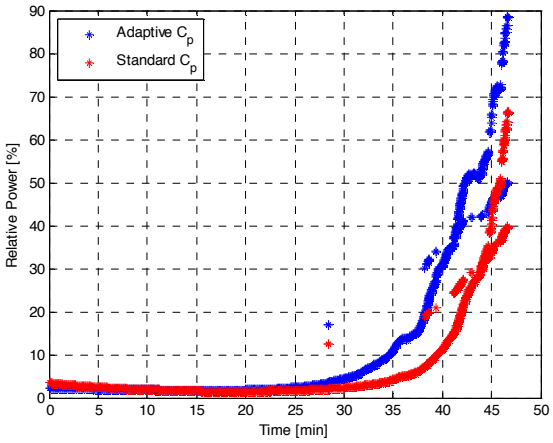


Fig. 6 Comparison of the simulations of power produced by the wind turbine for both adapted C_p and standard C_p.

This plot shows the power generated in the two cases of power references; the adapted one and the non-adapted one. It is clear to see that if the power controller is not adapted to

actual blades and C_p-table/curve quite allot of energy is actual lost. It is approximately 20 % on the maximal power. It is a further study to investigate how much the C_p-table varieties between wind turbines.

Using this scheme the torque reference curves can be adapted during the first couple of months during the commission of the wind turbine using only standard measurements present at the wind turbine. An alternative would be to place a wind measuring mast in front of each turbine which would be highly costly, and consequently almost never done in practice. Even though wind measurement equipments are used at a few test turbines to document the power production. This measurement could be used to adapt the power reference curve, and is some times done. It will, however, clearly increase the potential gain if every turbine is adapted to its own actual power reference curve which is possible using the proposed algorithm.

VI. Conclusions

Optimality of power produced by wind turbines depends on the correctness of a table of power coefficients. This table is normally provided by the manufactures of the wind turbine blades. These are measured or calculated for a limited number of blades in a production series but not for all produced blades. This means that the used table might be incorrect regarding the actual used one. This again leads to risk of non-optimal power production by the wind turbine. An unknown input observer designed in another paper is used to estimate a new C_p-table which fits the actual power coefficients of the blades of the wind turbine. This might require a number of iterations if the actual power coefficients are highly different from the assumed C_p-table. The optimal reference curve is found based on this newly computed C_p-table. A simulation is performed where the found optimal reference curve is compared with the initial assumed one. In this simulation the non-adapted power controlled wind turbine generates approximately 20% less power compared with the adaptive power controlled one.

VII. Refences

[1] K. Johnson, M. Pao, L.Y.and Balas, and L. Fingersh, "Control of variable-speed wind turbines - standard and adaptive techniques for maximizing energy capture," *IEEE Control Systems Magazine*, vol. 26, no. 3, pp. 71–81, June 2006.

[2] Y. Song, B. Dhinakaran, and X. Bao, "Variable speed control of wind turbines using nonlinear and adaptive algorithms," *Journal of Wind Engineering and Industrial Aerodynamics*, vol. 85, pp. 293–308, 2000.

[3] R. Sakamotol, T. Senjyul, T. Kinjo, N. Urasaki, and T. Funabashi, "Output power leveling of wind turbine generator by pitch angle control using adaptive control method," in *Proceedings of 2004 lnternational Conference on Power System Technology - POWERCON 2004*, Singapore, November 2004, pp. 834–839.

[4] A. Dadone and L. Dambrosio, "Estimator based adaptive fuzzy logic control technique for a wind turbine generator system," *Energy Conversion and Management*, vol. 44, pp. 135–153, 2003.

[5] P. Odgaard, C. Damgaard, and R. Nielsen, "On-line estimation of wind turbine power coefficients using unknown input observers," in *Proceedings of the 17th World Congress The International Federation of Automatic Control*, IFAC. Seoul, Korea: IFAC, July 2008, pp. 10646–10651.

[6] J. Chen and R. J. Patton, *Robust model-based fault diagnosis for dynamic systems*, 1st ed. Kluwer academic publishers, 1999.

[7] P. Odgaard and B. Mataji, "Observer based fault detection and moisture estimating in coal mill," *Control Engineering Practice*, vol. 16, no. 8, pp. 909–921, August 2008.

[8] P. Odgaard, C. Damgaard, and R. Nielsen, "Adaptive power control of wind turbines based on power coefficient estimations," in *Proceedings of European Wind Energy Conference 2008*. Brussels, Belgium: EWEA, April 2008, Obtained at: http://www.ewec2008proceedings.info/ewec2008/allfiles2/10_EWEC2008fullpaper.pdf

11 Critical fault detection, by measured current on electromechanical hydraulic valves

Critical fault detection, by measured current on electromechanical hydraulic valves

Lasse Skov
kk-electronic a/s
Ikast, Denmark
lasko@kk-electronic.com

Peter Fogh Odgaard
kk-electronic a/s
Viby J, Denmark
peodg@kk-electronic.com

Rasmus Nielsen
kk-electronic a/s
Ikast, Denmark
ranie@kk-electronic.com

***Abstract*— Fault detection in large scale industrial systems are of importance in order to detect and accommodate eventual faults and failures in the system. Such an industrial system is a hydraulic system which typically consists of a high number of control valves. It would be beneficial to detect eventually faults in these valves. In computer controlled systems these valves are often electromagnetic actuated. In this paper a model of an electromagnetic normally-closed hydraulic valve is completed. This model of the valve is used to test a scheme to detect blocked gliders in the electromagnetic actuated hydraulic valve. The scheme tested in this paper only detects when a valve is blocked. This is a very important feature to have in a system with many hydraulic valves, since it can take a lot of time to find the valve with the failure in such a system. In a situation like this it is critical to detect and isolate a blocked valve; however, it is not necessary to detect the fault at the instant it occurs. The algorithm used in this paper find 100% faulty cases, without any false positive detection.**

Keywords-Fault Detection, Hydraulic Valves

I. INTRODUCTION

In the industry valves are often used to control hydraulic systems. In a complex system with many hydraulic valves, it will be highly time consuming to localize a defect valve. Therefore it could be useful to have functionality in the control system to detect if a valve is moving or not as it should.

The valve used in this paper is at normally open, double blocking poppet-type solenoid valve. This valve is activated by a 26W coil and the magnetic will push the masses and close the valve. This armature is pushing on a small mass which again is pushing on the plunger. The plunger closes and opens for the hydraulic oil.

In order to have robust detection of faults in the valve in terms of a blocked plunger, the best solution is to use the coil current measurement as the detection signal. This current is measured in order to control the proportional valve.

This controller's actuation signal is in most case a pulse wide modulated signal, since this enables a design of a controller to reduce the current to what is necessary to hold the valve open. The control loop is formed with the coil current as measured output and the voltage over the coil as control signal.

Most of the published research on fault detection on hydraulic systems has been on the system level. Detection on specific valves has been rare. Some examples are as follows.

In [8] a neural network based scheme is used to identify the parameters in physical model of a hydraulic system. In [5] redundancy and logic based detection scheme are used to find the most likely "good" sensor signal and actuator drive. In [2] an Extended Kalman estimator is used to detect faults in a hydraulic actuator system with servo valves and pistons. A model-based fault detection scheme is presented in [4] which is used to detect faults in hydraulic brake system for automobiles. In [3] and [11] two different fault detection benchmark problem are presented dealing with the industrial problem of detection of faults in valves in process plant. A couple of examples on detection valve faults in case of process control systems are: stiction detection and control in [12], and fuzzy logic based detection in [6].

Examples on models of electrical actuated hydraulic valves can be seen in [14, 1, 16, 9, 15], and a more basic description of the physics of the valves can be seen in [7].

In [10] and unknown input observer is used in the same scenario this scheme can detect the faults as they occurs. In [13] a number of different detection schemes have been suggest to detect the fault in an electro activated hydraulic valve, one method has as well been implemented, a method based on differentiation has been implemented. The detection on the valve is not time critical, so it could be possible to record some data and make an analysis on the recorded dataset. The same data used in these experiments is described in this paper and the differentiation detection is described. Before the algorithm the model used for test is described in details, in order to provide information about the valves mechanical data, and the test setup.

A FFT analysis is performed on the fault free and faulty model output to compute the frequency range of the major difference between the two cases.

In Section II a model of the electro activated hydraulic valve is presented, and in Section III an experimental setup is presented. Validation of the model is presented in Section IV, and the proposed scheme is presented in Section V, followed by simulation results in Section VI. The paper is concluded in Section VIII.

II. Model of Electroactivated Hydraulic Valve

A. Systems Description

The system in question is an on/off electro-magnetic hydraulic valve (normally closed) in which the position is controlled by the voltage over the coil, $v_{\mathrm{sol}}(t)$, and only the coil current, $i(t)$, is measured and consequently provided as output. A current controller is form to control and position the valve glider. An illustration of the valve can be seen in Fig. 1

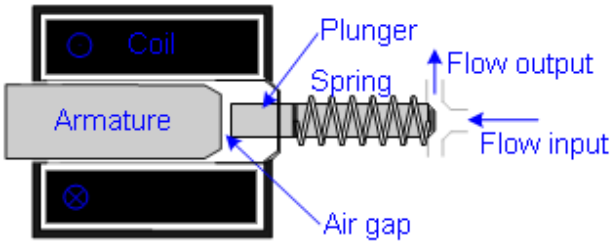


Fig. 1 An illustration of the physical layout of the on/off electro-magnetic hydraulic valve.

In the dynamic model of the electro-magnetic valve there are two deferent situations. Situation 1 is where the armature is open and the armature is moving to the plunger. Situation two is where the two masses are moving together closing the valve, see Fig. 2. The figure also defines the minimal, maximal and servo positions, d_{min}, d_{max}, d_{ser}.

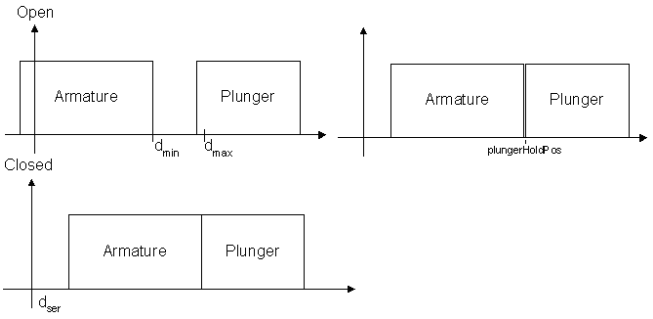


Fig. 2 An illustration of the two situations of the valve movements.

B. Model of Valve

In the section the dynamic equations for the valve will be described. The coil has N turns and a resistance R. The armature inside the coil has an area A, a mass m_a, and plunger sitting in the valve has a spring constant F_s, and the preloaded force F_{s0}. Area of the plunger is A_0, mass of it is m_p, and the pressure from the hydraulic supply is p_s. In the two situations the frictions constants are different. The first situation the friction constant is C_{v1}, and in the next situation the constant is C_{v2}.

The magnetic field, $B(t)$, can be found by

$$B(t) = \frac{1}{A}\int \frac{1}{N}(v_{sol}(t) - R \cdot i(t))dt, \tag{1}$$

Resulting in a differential equation

$$\dot{B}(t) = \frac{1}{A \cdot N}(v_{sol}(t) - R \cdot i(t)), \tag{2}$$

The magnetic force, $F_{\mathrm{sol}}(t)$ is

$$F_{sol}(t) = \frac{1}{2}\frac{A}{\mu_0} \cdot B(t)^2, \tag{3}$$

The current is

$$i(t) = \frac{1}{N}\left(\frac{B(t)}{\mu_0} \cdot x_{cur}(t)\right), \tag{4}$$

Where

$$x_{cur}(t) = d_{max} - x(t) + d_{ser} - d_{min} \tag{5}$$

The position $x(t)$ is found by two situations depending on its value. In (6)-,(7) the acceleration, $\ddot{x}(t)$, of the position $x(t)$ are found.

If $x(t) < x_{hold}$

$$\ddot{x}(t) = \frac{1}{m_a}\left(F_{sol}(t) - C_{v1}\dot{x}(t)\right), \tag{6}$$

If $x(t) \geq x_{hold}$

$$\ddot{x}(t) = \frac{(F_{sol}(t) - A_0P_s - K_sx(t) - F_{s0} - C_{v2}\dot{x}(t))}{m_a + m_p}, \tag{7}$$

in which x_{hold} is the meeting position of the plunger and armature.

This gives two nonlinear state space representations:

Situation 1

$$\begin{bmatrix}\dot{B}(t)\\ \ddot{x}(t)\\ \dot{x}(t)\end{bmatrix} = f_1\left(\begin{bmatrix}B(t)\\ \dot{x}(t)\\ x(t)\end{bmatrix}\right) + \begin{bmatrix}\frac{v_{sol}(t)}{A \cdot N}\\ 0\\ 0\end{bmatrix}, \tag{8}$$

Where

$$f_1\left(\begin{bmatrix}B(t)\\ \dot{x}(t)\\ x(t)\end{bmatrix}\right) = \begin{bmatrix}\frac{-R}{A \cdot N^2 \cdot \mu_o}B(t) \cdot x_{air}(t)\\ \frac{1}{m_a}\left(\frac{A}{2 \cdot \mu_0}B^2(t) - C_{v1}\dot{x}(t)\right)\\ \dot{x}(t)\end{bmatrix}, \tag{9}$$

Situation 2

$$\begin{bmatrix}\dot{B}(t)\\ \ddot{x}(t)\\ \dot{x}(t)\end{bmatrix} = f_2\left(\begin{bmatrix}B(t)\\ \dot{x}(t)\\ x(t)\end{bmatrix}\right) + \begin{bmatrix}\frac{v_{sol}(t)}{A \cdot N}\\ 0\\ 0\end{bmatrix}, \tag{10}$$

$$f_2\left(\begin{bmatrix} B(t) \\ \dot{x}(t) \\ x(t) \end{bmatrix}\right) = \begin{bmatrix} \frac{-R}{AN^2 \cdot \mu_o} B(t) x_{air}(t) \\ \frac{\left(\frac{A}{2\mu_0} B^2(t) - C_{v2}\dot{x}(t) - A_0 P_s - F_{s0} - K_s x(t)\right)}{m_a + m_p} \\ \dot{x}(t) \end{bmatrix}, \quad (11)$$

Where

$$x_{air}(t) = d_{max} - x(t) + d_{ser} - d_{min}, \quad (12)$$

The output equation is

$$i(t) = \frac{1}{N \cdot \mu_0} B(t) \cdot x_{air}(t), \quad (13)$$

The following parameters are used in the model:

R	A	A_0	N
31.0570Ω	$2.1517 \cdot 10^{-4} m^2$,	$9.021 \cdot 10^{-6} m^2$	2101.1
μ_0	m_a	m_p	p_s
$\pi \cdot 4 \cdot 10^{-7}$	$134.96 \cdot 10^{-3} kg$	$109.13 \cdot 10^{-3} kg$	$2 \cdot 10^6 Pa$
C_{v1}	C_{v2}	F_{s0}	K_s
$1.9656 \frac{N \cdot s}{m}$	$1.7353 \frac{N \cdot s}{m}$	$6.8954 N/m$	$23.512 N/m$
d_{max}	d_{min}	d_{ser}	χ
$7.2267 \cdot 10^{-3} m$	$5.9698 \cdot 10^{-3} m$	$2.3394 \cdot 10^{-3} m$	0.0863
x_{hold}			
$3.3765 \cdot 10^{-4} m$			

III. Experimental Setup

A system overview over the setup used in these experiments is shown in Fig. 3.

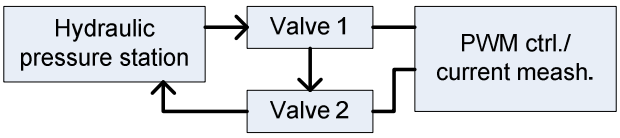


Fig. 3 – Block diagram over the experimental setup.

The hydraulic pressure station has a reservoir with hydraulic oil and an oil cooler included. The pump generates a pressure on 20[bar], and the valve used on the hydraulic station is normally-open, double blocking poppet-type solenoid valve.

The controller card has a Pulse Wide Modulation (PWM) output with a PI current regulator implemented. The power supply on the PWM output is a 24[V] supply. And it measure the current with the sample rate on 1[kHz].

IV. Experiments and Model Identification

Two experiments where preformed on the valves; one where the valve is blocked and another one where the valve is not. In both situations the input to the controller is a step on the current. The controller is a PI with anti windup.

K_p and T_i is tuned so the current controller not will reduce the current in the step. The following controller parameters are used: $K_p = 40, T_i = 5$.

The results of the experiments can be seen in Fig. 4, it is easy to see the difference between the blocked and non-blocked valve.

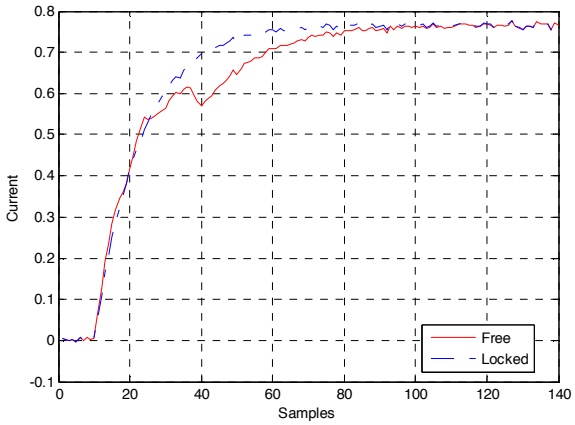


Fig. 4 Measured data from a situation with the same valve blocked and unblocked

With those data it is now possible to compare the measured data with the simulations using the model found for the valve in the previously section. After tuning of the valve parameters, on the non-blocked model the measured and simulated current is quite close, see Fig. 5. In case of the blocked valve the simulated current increased a bit faster than the measured current, see Fig. 6. Fortunately, this model error is not of large importance due to the large variation between blocked and non-blocked behavior

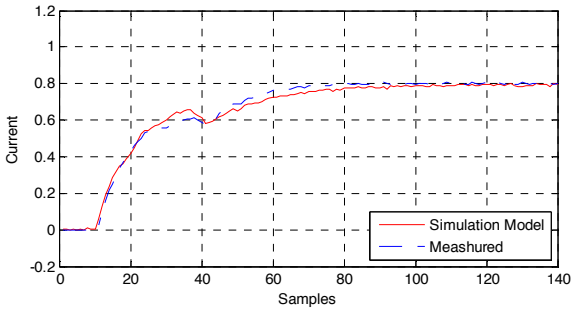


Fig. 5 Comparison on the model found and the measured data on the unblocked valve.

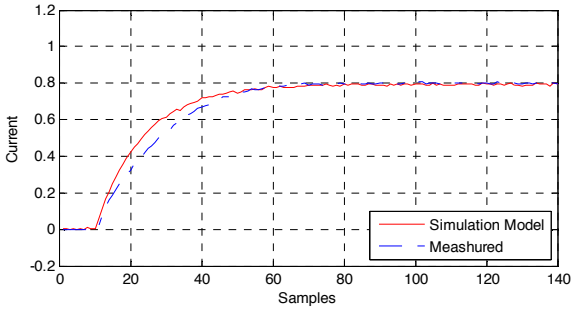


Fig. 6 Comparison between the model and the blocked valve.

Measurement noise is added to the simulated current as Gaussian noise with zero mean value and a std twice the measured one in the experimental setup $\sigma = 0.008$.

V. Proposed Scheme

In this paper scheme which is easy to implement and as well have shown strong performance is described.

It is a simple but effective scheme. It is based on a backward-differentiation of the current, which is used to find the catachrestic of the current, when the valve not is locked. This detection method is really useful in this scenario because the valve current contain much information about the armature and plungers position.

The backward-differentiation results in a high pass filter effect. On the measurements board there is a low pass filter. Combing these two gives a band pass filter. This input filter is a RC coupling witch transfer function there could be described by.

$$T(s) = \frac{\frac{1}{sC}}{R+\frac{1}{sC}} = \frac{\frac{1}{sC}}{sCR+1} = \frac{1}{1+sRC}, \tag{14}$$

in which: $R = 2.6 \cdot 10^3 \Omega, C = 2.2 \cdot 10^{-}9F.$

Transferring this function from the continues time domain to the discrete with using bilinear transformation with a sample frequency at 1[kHz]:

$$T(Z) = \frac{b_0 Z + b_1}{Z + a_1}, \tag{15}$$

where the coefficients for this transferred with bilinear approximation is, $b_0 = 0.984$, $b_1 = 0.984$, $a_1 = 0.984$. Implementation of the backward-differentiation is as following.

$$\alpha[n] = \frac{d}{dt} i[n] = \frac{i[n] - i[n-k]}{k \cdot \Delta}, \tag{16}$$

The transfer function of this is given by.

$$\alpha(Z) = \frac{1}{k \cdot \Delta} \cdot (I(Z) - Z^{-k} \cdot I(Z)), \tag{17}$$

$$\alpha(Z) = \frac{1}{k \cdot \Delta} \cdot (1 - Z^{-k}) \cdot I(Z), \tag{18}$$

$$\frac{\alpha(Z)}{I(Z)} = \frac{1}{k \cdot \Delta} \cdot \frac{Z^k - 1}{Z^k}, \tag{19}$$

The transfer function for the two systems will subsequently be folded to find the transfer function for the filter. And k is set to 2.

$$H_{bp}(Z) = \frac{\alpha(Z)}{I(Z)} \cdot T(Z), \tag{20}$$

$$H_{bp}(Z) = \frac{b_0 Z^3 + b_1 Z^2 - b_0 Z - b_1}{Z^3 + a_1 Z^2}, \tag{21}$$

The bode plot of this transfer function can be seen in Fig. 7. It is seen that this filter pass most of the energy in the frequency interval between 250[rad/s] and 2000[rad/s], which corresponds to (39.8[Hz] 318[Hz]).

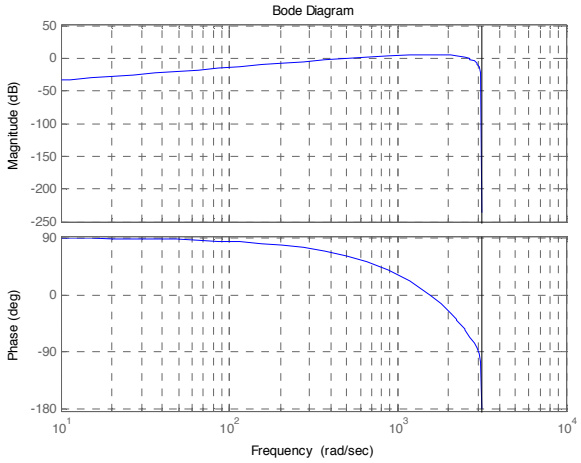


Fig. 7 Bode plot of the merged transfer function of the input filter and the detection filter.

This differentiation based detection filter works since the current increase much faster in case of a blocked valve compared to a non blocked valve. It reaches the steady state value after 60 samples compared to 80 samples for a non blocked valve.

After the backward-differentiation is calculated with $\Delta = 2$ the result $\alpha[n]$ is compared to the threshold, γ. The value of this variable is found by trail error method; in many different scenarios on the model. The most optimal threshold value for this valve is $\gamma = -5$.

A. *FFT analysis of faulty and fault free valve*

The characteristic in the current generated by coil can be computed with a Fast Fourier Transform (FFT) the movement of the mass inside the coil will give a specific frequency in the current.

To search this for frequency a windowed FFT is calculated, on the data generated from the model, to se if it is possible to find the characteristic frequency in the current. This analysis is executed in order to find the frequency range with the largest difference between the faulty and fault free case.

Fig. 8 shows the frequency response in the current when a valve is in normal operation. It is clear to see that most of the energy is located in frequencies in the interval between 40 and 50[Hz].

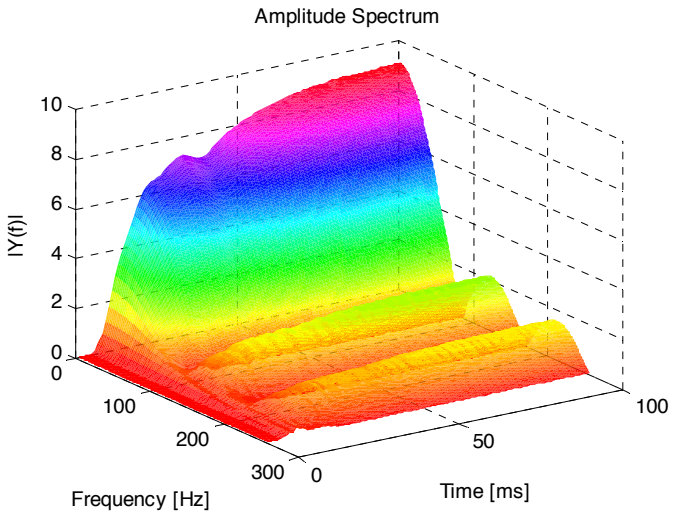


Fig. 8 FFT on data from fault free valve.

If a FFT based on same method is calculated when plunger is locked the calculations gives a result shown in Fig. 9.

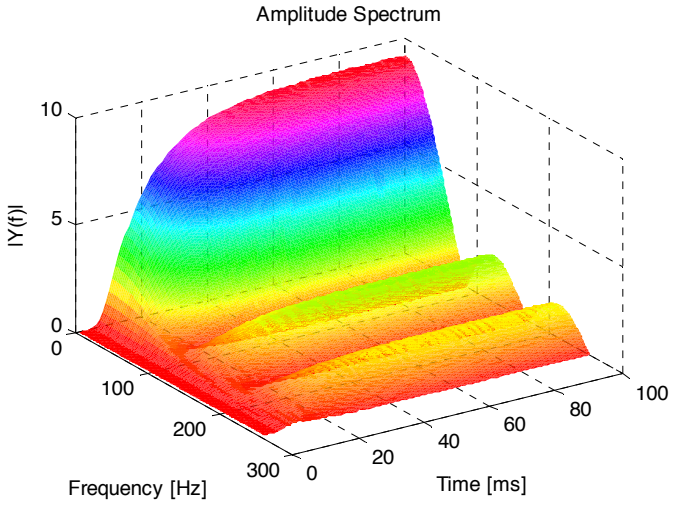


Fig. 9 FFT on data from a blocked valve.

In this plot it is simple to see there is a difference to Fig. 9. Again most of the energy is located between 40 and 50[Hz] but with a clearly difference sequence at these frequencies in time. This means that the frequency interval to look at for the detection filter should be between 40 and 50[Hz], and this interval the detection filter has a gain at around 4 db, meaning most of the energy of the error would pass through the detection filter. Consequently the detection filter is sensible towards the fault as it was suppose to.

FFT is in both examples calculated with at window size of 5 over 100 samples.

VI. Simulations

The simulations are performed using the backward-differentiation detector where 2744 situations are simulated, 1372 situations are fault free and 1372 has a fault. This fault is blocked value at a position between 0.6mm and 1.4mm.

The results of the simulations tests executed on the algorithm with the model give following results.

True Positive	100 %
False Negative	100 %
False Positive	0 %
True Negative	0 %

Table 1 - Tabel of preformed test on algorithm.

This algorithm detects the failure pretty good, beside it is quite simple to implement. The rate of correct detections are that high since the current in the blocked case increases dramatically faster than in the non-blocked case. Due to the model uncertainty, the simulation of the blocked valve might result in a much clearer difference between the blocked and non-blocked case, this means that in practice the results might a little bit less impressive in terms of true positive detection rates.

VII. Conclusions

In this paper a model of an electromechanically valve is found and used to test goodness of algorithm for detecting fouls on the valve in terms of blocked valves. The only measured data is the current and this is differentiated and from this it is easy to see if there is a critical fault on the valve. 1372 faults and 1372 non faults are simulated with measurement noise. With a correct threshold value the detector detect 100% true negative faults, and 100% true positive faults.

References

[1] Andrew Alleyne and Rui Liu. A simplified approach to force control for electro-hydraulic systems. *Control Engineering Practice*, 8(12):1347–1356, December 2000.

[2] L. An and N. Sepehri. Hydraulic actuator circuit fault detection using extended kalman filter. In N. Sepehri, editor, *Proc. American Control Conference the 2003*, volume 5, pages 4261–4266 vol.5, 2003.

[3] Michal Bartys, Ron Patton, Michal Syfert, Salvador de las Heras, and Joseba Quevedo. Introduction to the damadics actuator fdi benchmark study. *Control Engineering Practice*, 14(6):577–596, June 2006.

[4] Marcus Borner, Harald Straky, Thomas Weispfenning, and Rolf Isermann. Model based fault detection of vehicle suspension and hydraulic brake systems. *Mechatronics*, 12(8):999–1010, October 2002.

[5] Salem Haggag, Aristoteles Rosa, Kevin Huang, and Sabri Cetinkunt. Fault tolerant real time control system for steer-by-wire electro-hydraulic systems. *Mechatronics*, 17(2-3):129–142, March- April 2007.

[6] L.F. Mendonca, J.M.C. Sousa, and J.M.G. Sa da Costa. Fault isolation using fuzzy model-based observers. In *Proceedings of the Sixth IFAC Symposium on Fault Detection, Supervision and Safety of Technical Processes, 2006*, volume 6, pages 735–740, Beijing, China, August 2006.

[7] Herbert E. Merritt. *Hydraulic control systems*. John Wiley & Sons, Inc., - edition, 1967.

[8] Marco Muenchhof. Semiphysical models of a hydraulic servo axis for fault detection. In *Proc. American Control Conference ACC '07*, pages 1834–1839, New York, NY, USA, July 2007. ACC, IEEE.

[9] Huba Nemeth, Piroska Ailer, and Katalin M. Hangos. Nonlinear modelling and model verification of a single protection valve. *Periodica Polytechica Transportation Engineering*, 30(1-2):69–92, 2002.

[10] P. F. Odgaard, L. Skov, and R. Nielsen. Unknown input observer based detection scheme for faults in hydraulic valves. 2009.

[11] R. Patton. A benchmark study approach to fault diagnosis of industrial process control systems. In *The IEE Seminar on Control Loop Assessment and Diagnosis, 2005*, pages 61–79, June 2005.

[12] A. Singhal and T.I. Salsbury. A simple method for detecting valve stiction in oscillating control loops. *Journal of Process Control*, 15(4):371–382, June 2005.

[13] Lasse Skov. Ventil fejl detektion (fault detection in valves). B.sc(EE) Thesis AU-HIH, December 2007. In Danish.

[14] Jose R. Valdes, Miana Mario J., Jose L. Nunez, and Thomas Putz. Reduced order model for estimation of fluid flow and flow forces in hydraulic proportional valves. *Energy Conversion and Management*, In Press, Corrected Proof, 2008.

[15] Q.H. Yuan and P. Y. Li. Self-calibration of push-pull solenoid actuators in electrohydraulic valves. In *Proceedings of 2004 ASME International Mechanical Engineering Congress and RD&D Expo*, Anaheim, CA, USA, November 2004. ASME.

[16] Qinghui Yuan and Perry Y. Li. Using steady flow force for unstable valve design: Modeling and experiments. *Journal of Dynamic Systems, Measurement, and Control*, 127(3):451–462, September 2005.

12 Unknown Input Observer Based Detection Scheme for Faults in Hydraulic Valves

Unknown Input Observer Based Detection Scheme for Faults in Hydraulic Valves

Peter Fogh Odgaard
kk-electronic a/s
Viby J, Denmark
peodg@kk-electronic.com

Lasse Skov
kk-electronic a/s
Ikast, Denmark
lasko@kk-electronic.com

Rasmus Nielsen
kk-electronic a/s
Ikast, Denmark
ranie@kk-electronic.com

***Abstract*— In this paper an observer based scheme is proposed to detect blocked gliders in electromagnetic actuated hydraulic valves. Detection of blocked gliders in electromagnetic actuated hydraulic valves is of large importance in large hydraulic systems which contain a large number of control valves. If the glider position was measured this detection would be simple, however, in many cases only the coil current is measured. In case of a blocked glider, it can be viewed as the introduction of an extra force to keep the glider in position. This extra force can be viewed as an unknown input, and can be estimated by the use of an unknown input observer. Using this estimated fault signal gives correct fault detection in case of 96.6% of 1500 faulty cases and only 4% in case of no faults, these numbers are found using Monte Carlo simulations. The detection time is as fast as the step response of the current changes in the valve.**

Keywords-Fault Detection, Hydraulic Valves, Unknown Input Observers

I. INTRODUCTION

In the industry hydraulic systems valves are often used to control the systems. In case of a complex system with many hydraulic valves, it will take a lot of time to localize a defect valve in the system, and sometime a defect valve could result in large damage of the system. To find this defect on the valve before it will damage on the system would be very helpful. The system operator could as well be informed that this valve is going to be damaged, so it could be replaced with a new one before it damages the system.

Most of the published research on fault detection on hydraulic systems has been on the system level. Detection on specific valves has been rare. Some examples are as follows.

In [1] a model of a hydraulic servo system is made by using a neural network to identify the parameters in physical based model of the hydraulic system with a neural network based model on this structure. A multi model approach for detection of faults in a hydraulic servo axis can be seen in [2]. In [3] redundancy and logic based detection scheme are used to find the most likely "good" sensor signal and actuator drive. In [4] an Extended Kalman estimator is used to detect faults in a hydraulic actuator system with servo valves and pistons. A model-based fault detection scheme is presented in [5] which is used to detect faults in hydraulic brake system for automobiles. In [6] a fault detection benchmark problem is presented dealing with the industrial problem of detection of faults in valves in process plant. Signal processing on the coil current is often used for fault detection in process plant valves. In [7] a number of different detection schemes have been suggest to detect the fault in a electro activated hydraulic valve, one method has as well been implemented; a method based on differentiation has been implemented. Examples on models of electrical actuated hydraulic valves can be seen in [8, 9, 10, 11, 12], and a more basic description of the physics of the valves can be seen in [13].

An alternative approach would be to use an observer based scheme such that faults can be detected based on both position measurements as well as control actions. In order to deal with model uncertainties, as well, an unknown input observer, see [14], is applied to detect the fault.

The valve in question is controlled by a current controller and consequently it would be beneficial to include an observer in the fault detection scheme, since the observer can combine information in both control signals and system output. A blocked glider in the valve could be modeled as a force balance with an additional force keeping the glider in a specific position whatsoever forces the controllers puts on the glider. In [15] and [16] an unknown input observer is used to detect faults and or missing input signals to systems, in the first the observer is used to estimate faults in coal mills and in the second paper it is used to estimate power coefficients and wind speeds in an application of a wind turbine. The same approach is applied to the problem of detecting blocked gliders in this specific hydraulic valve.

In this paper the system description and the model of the valve is first described, followed by the applied scheme and experiments. Finally a conclusion is drawn.

II. MODEL OF ELECTROACTIVATED HYDRAULIC VALVE

A. Systems Description

The system in question is an on/off electro-magnetic hydraulic valve (normally closed) in which the position is controlled by the voltage over the coil, $v_{\mathrm{sol}}(t)$, and only the coil

current, $i(t)$, is measured and consequently provided as output. A current controller is form to control and position the valve glider. An illustration of the valve can be seen in Fig. 1

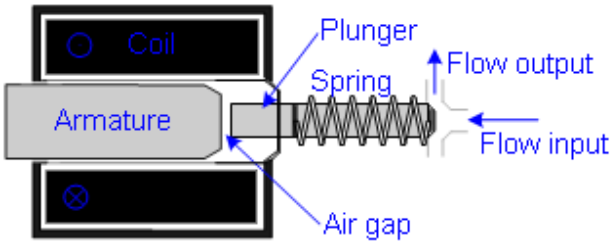


Fig. 1 An illustration of the physical layout of the on/off electro-magnetic hydraulic valve.

In the dynamic model of the electro-magnetic valve there are two deferent situations. Situation 1 is where the armature is open and the armature is moving to the plunger. Situation two is where the two masses are moving together closing the valve, see Fig. 2. The figure also defines the minimal, maximal and servo positions, $d_{\min}$, $d_{\max}$, d_{ser}.

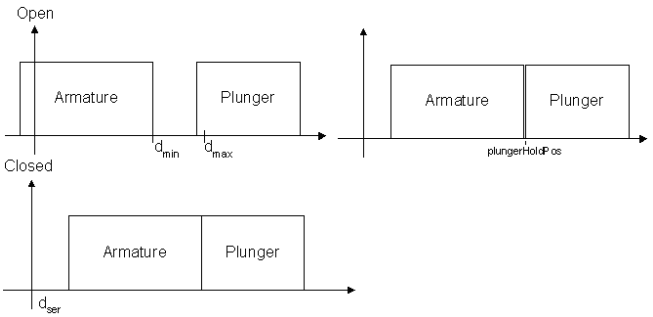


Fig. 2 An illustration of the two situations of the valve movements.

B. *Model of Valve*

In the section the dynamic equations for the valve will be described. The coil has N turns and a resistance R. The armature inside the coil has an area A, a mass m_a, and plunger sitting in the valve has a spring constant F_s, and the preloaded force F_{s0}. Area of the plunger is A_0, mass of it is m_p, and the pressure from the hydraulic supply is p_s. In the two situations the frictions constants are different. The first situation the friction constant is C_{v1}, and in the next situation the constant is C_{v2}.

The magnetic field, $B(t)$, can be found by

$$B(t) = \frac{1}{A}\int \frac{1}{N}(v_{sol}(t) - R\cdot i(t))dt, \qquad (1)$$

Resulting in a differential equation

$$\dot{B}(t) = \frac{1}{A\cdot N}(v_{sol}(t) - R\cdot i(t)), \qquad (2)$$

The magnetic force, $F_{\text{sol}}(t)$ is

$$F_{sol}(t) = \frac{1}{2}\frac{A}{\mu_0}\cdot B(t)^2, \qquad (3)$$

The current is

$$i(t) = \frac{1}{N}\left(\frac{B(t)}{\mu_0}\cdot x_{cur}(t)\right), \qquad (4)$$

Where

$$x_{cur}(t) = d_{max} - x(t) + d_{ser} - d_{min} \qquad (5)$$

The position $x(t)$ is found by two situations depending on its value. In (6)-,(7) the acceleration, $\ddot{x}(t)$, of the position $x(t)$ are found.

If $x(t) < x_{hold}$

$$\ddot{x}(t) = \frac{1}{m_a}(F_{sol}(t) - C_{v1}\dot{x}(t)), \qquad (6)$$

If $x(t) \geq x_{hold}$

$$\ddot{x}(t) = \frac{(F_{sol}(t) - A_0P_s - K_sx(t) - F_{s0} - C_{v2}\dot{x}(t))}{m_a + m_p}, \qquad (7)$$

in which x_{hold} is the meeting position of the plunger and armature.

This gives two nonlinear state space representations:

Situation 1

$$\begin{bmatrix}\dot{B}(t)\\ \ddot{x}(t)\\ \dot{x}(t)\end{bmatrix} = f_1\left(\begin{bmatrix}B(t)\\ \dot{x}(t)\\ x(t)\end{bmatrix}\right) + \begin{bmatrix}\frac{v_{sol}(t)}{A\cdot N}\\ 0\\ 0\end{bmatrix}, \qquad (8)$$

Where

$$f_1\left(\begin{bmatrix}B(t)\\ \dot{x}(t)\\ x(t)\end{bmatrix}\right) = \begin{bmatrix}\frac{-R}{A\cdot N^2\cdot\mu_o}B(t)\cdot x_{air}(t)\\ \frac{1}{m_a}\left(\frac{A}{2\cdot\mu_0}B^2(t) - C_{v1}\dot{x}(t)\right)\\ \dot{x}(t)\end{bmatrix}, \qquad (9)$$

Situation 2

$$\begin{bmatrix}\dot{B}(t)\\ \ddot{x}(t)\\ \dot{x}(t)\end{bmatrix} = f_2\left(\begin{bmatrix}B(t)\\ \dot{x}(t)\\ x(t)\end{bmatrix}\right) + \begin{bmatrix}\frac{v_{sol}(t)}{A\cdot N}\\ 0\\ 0\end{bmatrix}, \qquad (10)$$

$$f_2\left(\begin{bmatrix}B(t)\\ \dot{x}(t)\\ x(t)\end{bmatrix}\right) = \begin{bmatrix}\frac{-R}{AN^2\cdot\mu_o}B(t)x_{air}(t)\\ \frac{\left(\frac{A}{2\mu_0}B^2(t) - C_{v2}\dot{x}(t) - A_0P_s - F_{s0} - K_sx(t)\right)}{m_a+m_p}\\ \dot{x}(t)\end{bmatrix}, \qquad (11)$$

Where

$$x_{air}(t) = d_{max} - x(t) + d_{ser} - d_{min}, \qquad (12)$$

The output equation is

$$i(t) = \frac{1}{N\cdot\mu_0}B(t)\cdot x_{air}(t), \qquad (13)$$

The following parameters are used in the model:

R	A	A_0	N
31.0570Ω	$2.1517 \cdot 10^{-4}m^2$,	$9.021 \cdot 10^{-6}m^2$	2101.1
μ_0	m_a	m_p	p_s
$\pi \cdot 4 \cdot 10^{-7}$	$134.96 \cdot 10^{-3}kg$	$109.13 \cdot 10^{-3}kg$	$2 \cdot 10^{6}Pa$
C_{v1}	C_{v2}	F_{s0}	K_s
$1.9656\frac{N \cdot s}{m}$	$1.7353\frac{N \cdot s}{m}$	$6.8954N/m$	$23.512N/m$
d_{max}	d_{min}	d_{ser}	χ
$7.2267 \cdot 10^{-3}m$	$5.9698 \cdot 10^{-3}m$	$2.3394 \cdot 10^{-3}m$	0.0863
x_{hold}			
$3.3765 \cdot 10^{-4}m$			

In order to simplify the used observer only one linear model of the valve is used to design the observer. The point of operation is found in situation 2, causing a large model uncertainty then describing situation 1. The linear model can be seen in (14)-(17).

$$\begin{bmatrix} \dot{B}(t) \\ \ddot{x}(t) \\ \dot{x}(t) \end{bmatrix} = \mathbf{A} \cdot \begin{bmatrix} B(t) \\ \dot{x}(t) \\ x(t) \end{bmatrix} + \mathbf{B}v_{sol}(t), \tag{14}$$

$$i(t) = \begin{bmatrix} \frac{x_{air,0}}{N \cdot \mu_0} & 0 & \frac{B_0}{N \cdot \mu_0} \end{bmatrix} \cdot \begin{bmatrix} B(t) \\ \dot{x}(t) \\ x(t) \end{bmatrix}, \tag{15}$$

Where

$$\mathbf{A} = \begin{bmatrix} \frac{-R}{A \cdot N^2 \cdot \mu_0} \cdot x_{air,0} & 0 & \frac{-R}{A \cdot N^2 \cdot \mu_0} B_0 \\ \frac{A}{(m_a + m_p) \cdot \mu_0} B_0 & \frac{-C_{v2}}{m_a + m_p} & \frac{-K_s}{m_a + m_p} \\ 0 & 1 & 0 \end{bmatrix}, \tag{16}$$

$$\mathbf{B} = \begin{bmatrix} \frac{1}{A \cdot N} \\ 0 \\ 0 \end{bmatrix}, \tag{17}$$

$x_0 = 3.38 \cdot 10^{-4}m, x_{cur} = 3.3 \cdot 10^{-3}m,$

$B_0 = 0.0015,\ v_{sol,0} = 0.2012V, i_0 = 0.0020A.$

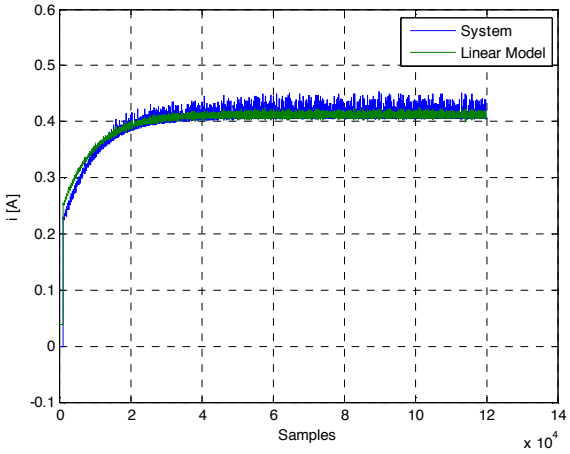


Fig. 3 Comparison of linear model output with system output

In Fig. 3 the linear model is compared with the system output, from which it can be seen that the linear model fits the system output quite well. The model is following discretized with a sample frequency at 1 kHz.

3. Proposed Schemes

The idea is to add an additional force to the motion model in,(7) and the second row in the state space model ,(11). Now take the linearized model (14)-(17) and introduce an unknown input into it representing a force keeping the glider in a blocked position. Subsequently an Unknown Input Observer is designed to estimate the unknown inputs. Consequently, in order to estimate the unknown input a state represent of the "fault" is introduced in the model, an example of this approach can be found in [15].

The introduced state represents the fast dynamics of the fault by a high pass filter.

$$\begin{bmatrix} \dot{B}(t) \\ \ddot{x}(t) \\ \dot{x}(t) \\ \dot{x}_f(t) \end{bmatrix} = \mathbf{A}_m \begin{bmatrix} B(t) \\ \dot{x}(t) \\ x(t) \\ x_f(t) \end{bmatrix} + \mathbf{B}_m v_{sol}(t) + \mathbf{E}_m f_f(t), \tag{18}$$

$$i(t) = \begin{bmatrix} \frac{x_{air,0}}{N \cdot \mu_0} & 0 & \frac{B_0}{N \cdot \mu_0} & 0 \end{bmatrix} \cdot \begin{bmatrix} B(t) \\ \dot{x}(t) \\ x(t) \\ x_f(t) \end{bmatrix}, \tag{19}$$

where: $f_f(t)$ is a unknown input representing the force needed to keep the gilder in the blocked position, $x_f(t)$ is the state representing the fault. The new state matrix is defined as $\mathbf{A}_m$ as, in which z_f is the zero of the high pass filter and p_f is the pole of the same filter.

$$\mathbf{A}_m = \begin{bmatrix} \frac{-R}{A \cdot N^2 \cdot \mu_0} \cdot x_{air,0} & 0 & \frac{-R}{A \cdot N^2 \cdot \mu_0} B_0 & 0 \\ \frac{A}{(m_a + m_p) \cdot \mu_0} B_0 & \frac{-C_{v2}}{m_a + m_p} & \frac{-K_s}{m_a + m_p} & z_f - p_f \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & -p_f \end{bmatrix}, \tag{20}$$

The $\mathbf{B}_m$ matrix has been extended with an extra zero.

$$\mathbf{B}_m = \begin{bmatrix} \frac{1}{A \cdot N} \\ 0 \\ 0 \\ 0 \end{bmatrix}, \quad (21)$$

The $\mathbf{E}_m$ matrix is defined as

$$\mathbf{E}_m = \begin{bmatrix} 0.1 \\ 0.1 \\ 0.1 \\ e_4 \end{bmatrix}, \quad (22)$$

The first three elements in $\mathbf{E}_m$ matrix is set to 0.1 in order to introduce some robustness regarding model uncertainties, and e_4 the parameter relating to the fault estimate will be found in the design section.

This linear model is subsequently discretized.

Since the system in mind contains some unknown inputs the idea is to use an unknown input observer in its optimal version, see (Chen & Patton, 1999). The structure is:

$$\begin{aligned} \mathbf{z}[n+1] =& \mathbf{F}_{n+1}\mathbf{z}[n] + \mathbf{T}_{n+1}\mathbf{B}_n\mathbf{u}[n] + \mathbf{K}_{n+1}\mathbf{y}[n], \\ \hat{\mathbf{x}}[n+1] =& \mathbf{z}[n+1] + \mathbf{H}_{n+1}\mathbf{y}[n+1], \end{aligned} \quad (23)$$

where $\mathbf{F}_{n+1}$, $\mathbf{T}_{n+1}$, $\mathbf{K}_{n+1}$ and $\mathbf{H}_{n+1}$ are matrices designed to achieve decoupling from the unknown input and as well obtain an optimal observer. $\hat{\mathbf{x}}[n]$ is a vector of the states of the extended model. The matrices in the unknown input observer are found using the following equation see (24)-(32).

$$\mathbf{E}_n = \mathbf{H}_{n+1}\mathbf{C}_{n+1}\mathbf{E}_n, \quad (24)$$

$$\mathbf{T}_{n+1} = \mathbf{I} + \mathbf{H}_{n+1}\mathbf{C}_{n+1}, \quad (25)$$

$$\mathbf{F}_{n+1} = \mathbf{A}_n - \mathbf{H}_{n+1}\mathbf{C}_{n+1}\mathbf{A}_n - \mathbf{K}^1_{n+1}\mathbf{C}_n, \quad (26)$$

$$\mathbf{K}^2_{n+1} = \mathbf{F}_{n+1}\mathbf{H}_n, \quad (27)$$

$$\mathbf{K}^1_{n+1} = \mathbf{A}^1_{n+1}\mathbf{P}_n\mathbf{C}^T_{n+1}\left(\mathbf{C}_n\mathbf{P}_n\mathbf{C}^T_n + \mathbf{R}_n\right)^{-1}, \quad (28)$$

$$\mathbf{A}^1_{n+1} = \mathbf{A}_n - \mathbf{H}\mathbf{C}_{n+1}\mathbf{A}_n, \quad (29)$$

$$\begin{aligned} \mathbf{P}_{n+1} =& \mathbf{A}^1_{n+1}\mathbf{P}^1_{n+1}\left(\mathbf{A}^1_{n+1}\right)^T + \mathbf{T}_{n+1}\mathbf{Q}_n\mathbf{T}^T_{n+1} \\ &+ \mathbf{H}_{n+1}\mathbf{R}_{n+1}\mathbf{H}^T_{n+1}, \end{aligned} \quad (30)$$

$$\mathbf{P}^1_{n+1} = \mathbf{P}_n - \mathbf{K}^1_{n+1}\mathbf{C}_n\mathbf{P}_n\left(\mathbf{A}^1_{n+1}\right)^T, \quad (31)$$

$$\mathbf{H}_{n+1} = \mathbf{E}_n\left(\mathbf{C}_n\mathbf{E}_n\right)^+, \quad (32)$$

The observer design procedure can be described by the following algorithm:

1) Set Initial values:

$\mathbf{P}_0 = \mathbf{P}(0)$,

$\mathbf{z}[0] = \mathbf{x}[0] - \mathbf{C}_0\mathbf{E}_0\left(\mathbf{C}_0\mathbf{E}_0\right)^+\mathbf{y}[0]$,

$\mathbf{H}_0 = \mathbf{0}$,

$k = 0$,

2) Compute $\mathbf{H}_{n+1}$ using (32).

3) Compute $\mathbf{K}^1_{n+1}$ and $\mathbf{P}^1_{n+1}$ using (28) and (31).

4) Compute $\mathbf{T}_{n+1}$, $\mathbf{F}_{n+1}$, $\mathbf{K}^2_{n+1}$ and $\mathbf{K}_{n+1}$ by (25), (26), (27) and $\mathbf{K}_{n+1} = \mathbf{K}^1_{n+1} + \mathbf{K}^2_{n+1}$.

5) Compute the state estimate $\hat{\mathbf{x}}[n+1]$ and $\mathbf{z}[+1]$ using (23).

6) Compute $\mathbf{P}_{n+1}$ using (30) and (31)..

7) Set $k = k + 1$ go to step 2).

3.1 Design of detection scheme

The observer and detection scheme is designed on data from the non-linear simulation model simulating a fault free situation and a situation with a blocked valve at 0.6mm. The two variance matrices in the observer design $\mathbf{R}$ and $\mathbf{Q}$ are found by experiments in order to optimize the correct fault detection rate and detection time, resulting in the following values:

$$R = 1000, \quad Q = 10^4 \cdot \mathbf{I} \quad (33)$$

$$e_4 = 1, \quad z_f = 2.83 \cdot 10^{-2}, \quad p_f = 2.89 \cdot 10^{-2} \quad (34)$$

Additional a threshold γ is used determine if a fault is present. If $x_f[n] > \gamma$ a fault is detected and no faults are detected if $x_f[n] \leq \gamma$. By experiments $\gamma = 0.15$ is found as the lowest threshold value which detects the faults without too many false positive detections.

The detection scheme applied to data from a blocked valve can be seen in Fig. 4. It can be seen that the fault is detected a approximately the same time as the step response has reached its steady state value, see Fig. 3.

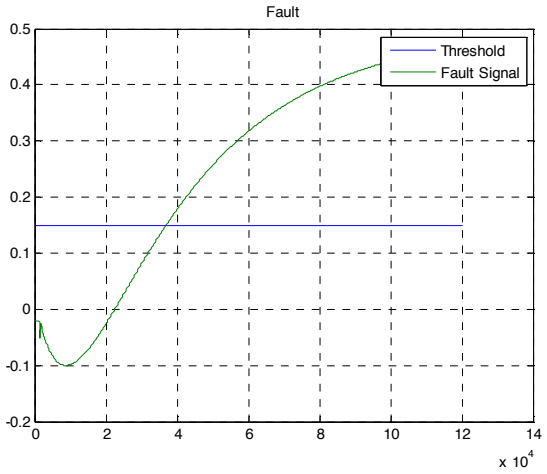


Fig. 4 Illustration of the fault estimate and threshold in case of a fault – a blocked value.

In Fig. 5 the fault estimate is compared with the threshold in case of a non blocked valve and consequently a fault is not detected.

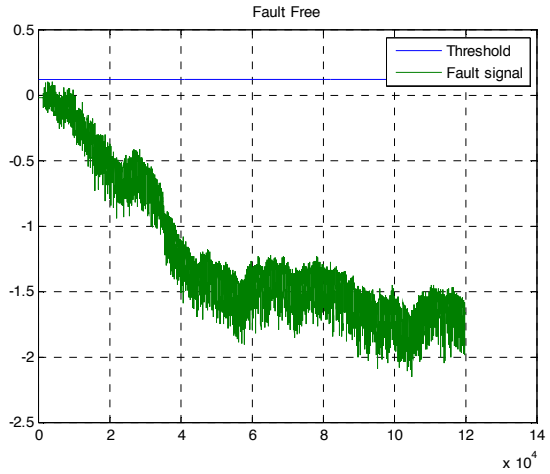


Fig. 5 Illustration of the fault estimate and threshold in case of no faults.

Experiments

The experiments are performed using Monte Carlo simulations on the nonlinear model where 3000 situations are simulated, 1500 situations are fault free and 1500 has a blocked value at a position between 0.6mm and 1.4mm. Measurement noise is added to the simulated current as Gaussian noise with zero mean value and a std (standard deviation) twice the measured one in the experimental setup $\sigma = 0.008$.

The results are shown in Table 1, where it is seen that 96.6% of the faults are correctly detected and only 4% of the non-faults are falsely detected as faults. It is, consequently seen that this proposed observer based detection scheme has a good performance of detecting the faults in the valve even though the presence of the large model uncertainties, which are tested by applying the observer on data from non-linear model some different model parameters. These tests are not a part of this paper, since they are on going, but they indicate some good properties regarding handling model uncertainties.

True Positive	96.6%
False Negative	3.4%
False Positive	4.0%
True Negative	96.0%

Table 1: Results of the simulation with 1500 faults and 1500 non-faults.

Other methods not focusing on fast detections can obtain 100 correct fault detections, which can be match by the proposed algorithm if detection time is not considered.

III. Conclusion

In this paper an unknown input observer based scheme is proposed to detect faults in hydraulic valves in terms of blocked gliders. The only measurement assumed is the current through the valve coil. The current is controlled to follow a step reference. 1500 faults and 1500 non-faults are simulated with measurement noise and the detection scheme detects in must cases the fault at the time the current reaches its steady state value. The detection scheme detects 96.6% "true positive faults" and 96% "true negative faults" (correctly detected non-faults).

IV. References

[1] M. Muenchhof, "Semiphysical models of a hydraulic servo axis for fault detection," in *Proc. American Control Conference ACC '07*, ACC. New York, NY, USA: IEEE, July 2007, pp. 1834–1839.

[2] ——, "Multi-model based fault detection and diagnosis of a hydraulic servo axis," in *Proceedings of the 17th World Congress The International Federation of Automatic Control*. Seoul, South Korea: IFAC, July 2008, p. 13117.

[3] S. Haggag, A. Rosa, K. Huang, and S. Cetinkunt, "Fault tolerant real time control system for steer-by-wire electro-hydraulic systems," *Mechatronics*, vol. 17, no. 2-3, pp. 129–142, March- April 2007.

[4] L. An and N. Sepehri, "Hydraulic actuator circuit fault detection using extended kalman filter," in *Proc. American Control Conference the 2003*, N. Sepehri, Ed., vol. 5, 2003, pp. 4261–4266 vol.5.

[5] M. Borner, H. Straky, T. Weispfenning, and R. Isermann, "Model based fault detection of vehicle suspension and hydraulic brake systems," *Mechatronics*, vol. 12, no. 8, pp. 999–1010, October 2002.

[6] M. Bartys, R. Patton, M. Syfert, S. de las Heras, and J. Quevedo, "Introduction to the damadics actuator fdi benchmark study," *Control Engineering Practice*, vol. 14, no. 6, pp. 577–596, June 2006.

[7] L. Skov, "Ventil fejl detektion (fault detection in valves)," Ba.EE Thesis AU-HIH, December 2007, in Danish.

[8] J. R. Valdes, M. M. J., J. L. Nunez, and T. Putz, "Reduced order model for estimation of fluid flow and flow forces in hydraulic proportional valves," *Energy Conversion and Management*, vol. In Press, Corrected Proof, 2008.

[9] A. Alleyne and R. Liu, "A simplified approach to force control for electro-hydraulic systems," *Control Engineering Practice*, vol. 8, no. 12, pp. 1347–1356, December 2000.

[10] Q. Yuan and P. Y. Li, "Using steady flow force for unstable valve design: Modeling and experiments," *Journal of Dynamic Systems, Measurement, and Control*, vol. 127, no. 3, pp. 451–462, September 2005.

[11] H. Nemeth, P. Ailer, and K. M. Hangos, "Nonlinear modelling and model verification of a single protection valve," *Periodica Polytechica Transportation Engineering*, vol. 30, no. 1-2, pp. 69–92, 2002.

[12] Q. Yuan and P. Y. Li, "Self-calibration of push-pull solenoid actuators in electrohydraulic valves," in *Proceedings of 2004 ASME International Mechanical Engineering Congress and RD&D Expo*. Anaheim, CA, USA: ASME, November 2004.

[13] H. E. Merritt, *Hydraulic control systems*, - ed. John Wiley & Sons, Inc., 1967.

[14] J. Chen and R. J. Patton, *Robust model-based fault diagnosis for dynamic systems*, 1st ed. Kluwer academic publishers, 1999.

[15] P. Odgaard and B. Mataji, "Observer based fault detection and moisture estimating in coal mill," *Control Engineering Practice*, vol. 16, no. 8, pp. 909–921, August 2008.

[16] P. Odgaard, C. Damgaard, and R. Nielsen, "On-line estimation of wind turbine power coefficients using unknown input observers," in *Proceedings of the 17th World Congress The International Federation of Automatic Control*, IFAC. Seoul, Korea: IFAC, July 2008, pp. 10646–10651.

13 A Simulation Based Investigation of Interactions between VVA and Idle Control for SI Engines

A Simulation Based Investigation of Interactions between VVA and Idle Control for SI Engines

Antonio Palma
and Ferdinando De Cristofaro
Systems and Controls, Gasoline EMS,
Elasis S.C.p.A. Via ex Aeroporto snc,
80038, Pomigliano d'Arco (NA) Italy.
Email: antonio.palma@fptpowertrain.elasis.it
ferdinando.decristofaro@fptpowertrain.elasis.it

Angelo Palladino
and Giovanni Fiengo
Dipartimento di Ingegneria,
Università degli Studi del Sannio,
Piazza Roma 21, 82100 Benevento, Italy.
Email: angelo.palladino@unisannio.it,
gifiengo@unisannio.it

Alessandra Guzzo
Università della Calabria
Rende (CS), Italy
Email: guzzo.alessandra@gmail.com

***Abstract*—Automotive idle speed control is a critical issue in engine control fields. Essentially it is a highly nonlinear and time-varying problem. Its performance has a significant impact on fuel economy and emission levels. In this paper, the authors present a complete and coherent engine model, aimed at the challenging purpose of the analysis of the interaction between the idle speed control and variable valve actuation system. The model is based on an innovative approach for engine dynamics conceived mainly on the analogy with electric systems. Firstly, the behavior of a relatively simple and well known control, named Mid-Ranging scheme has been tested on a complete "in-cylinder" engine model, after a modified control is proposed, in order to test the modern engine performances improvement due to the synergy between variable valve actuation devices and idle speed control strategy.**

Keyword: Modeling; Automotive system; SI engine model, VVA control.

I. Introduction

The evolution of engine control from mechanical to electronic strategy, and from open loop to closed loop control method, is a consequence of the hardware and software development. The use of electronics improves sensing accuracy, actuation capability, and flexibility in control law design. This results in adopting complex control strategy based on dynamic models purposely designed. The paper deals with the Idle Speed Control (ISC) and Variable Valve Actuation (VVA) systems based on an engine model formalized mainly on the analogy with electric systems.

The objective is to analyze the interaction between the engine idle speed and the VVA application. Load torque disturbances, such as air conditioning, power steering and VVA different commands can cause fluctuation on engine speed around the idle speed reference. Moreover, vehicle aging, variation in fuel efficiency, emissions and automotive vibrations can also effect the idle speed set point.

Several works are presented in literature on these topics. In (1), a review of various dynamic control technologies successfully applied to ISC systems is presented. The automotive ISC is a multi-objective and multi-variable control issue. The selection of target idle speed corresponds to a tradeoff among fuel consumption, idle operation stability, and emission levels (2). The idle speed control mechanism of different engine types varies. For a typical Port Fuel Injection (PFI) engine, the ISC system uses a controlled bypass valve to regulate the air flow rate around a closed throttle, using a solenoid, stepper motor, or controlled duty-cycle valve. The throttle bypass valve is used to adjust the intake manifold air flow around the primary throttle plate at idle. Its secondary function is to prevent stalling and produce smooth throttle tip in and tip out by providing extra air during idling and acting as an electronic dashpot during sudden deceleration (3).

Moreover, the presence of the VVA can cause uncertainty in the air mass flow estimation. In fact the different profiles of the VVA, full lift, early closing, late opening and multi lift, determinate a variable air mass flow depending by VVA control, as described in details in the following. So it is critical to model such system, aimed at analyzing the phenomenon and at designing suitable control strategies to compensate this uncertainty (4). The main contribution is the air-charge prediction for a Spark Ignition Internal Combustion Engine (SI-ICE) equipped with VVA, necessary to verify and validate control strategies, such as air-fuel regulation, idle speed control and torque generation (5). In particular, an accurate estimation of air flowing into the cylinders is the key point to determinate the engine fueling rate in order to control the air-fuel ratio during the combustion (6) and, consequently, to optimize the performance of the three way catalytic converter.

Generally, in order to overcome this kind of problem, mathematical models are used both to analyze the behavior of the system and, eventually, to design advanced control strategies. Among others, finite models of the combustion chamber allow to calculate all in-cylinder variables with a local resolution. The drawback is the long computing time. Vice versa, MVEMs (Mean Value Engine Models) simulate the dominant physical effects neglecting the fast dynamics, resulting in a lower computing time.

In this context, the paper presents an SI engine model equipped with a VVA system aimed at controlling the engine idle speed through two different control strategies, a multivariable mid-ranging algorithms and a new solution for the strategies to reduce the control complexity. The proposed

model is obtained exploiting an innovative approach based on the analogy of the engine components with electrical circuits (7). The robustness of the model is tested comparing some measurable variables with experimental data, while the analysis on the proposed solution is based only on simulation results due to lack of measurements.

II. Engine Model

In this paper, the engine model used to analyze the interaction between idle control and VVA system is derived by (7). In the following, the utilized modeling approach is briefly described. The authors, starting from a simple analogy with electrical systems, have obtained an engine description similar to an electrical circuit, with all the useful consequences in term of existence and numerical availability of the solution (8). The advantages are in the specific comparison that is founded between the engine components and variables (as throttle valve, cylinder, inertial flows) with electrical counterparts (current, voltage, resistance). In the following, the utilized modeling approach is briefly described. For details and equations see (7).

The engine is seen as an array of cylinders, having common connections with an intake and an exhaust manifold. The connections are regulated by valves opening. It is then possible to distinguish separate subsystems interconnected each others, such as the intake manifold equipped with throttle valve, the exhaust manifold and cylinders. From the phenomenological point of view, the elements composing the engine can be classified in the following categories: volumes, orifices, inertial effects and combustion.

The intake and exhaust manifolds and cylinders are grouped respectively as constant and variable volumes.

The orifices are responsible of the pressure drops along the gas path. They are modeled as variable resistances causing equivalent voltage drops.The size of the orifice. and consequently the resistance, is variable and regulated by valve opening, as throttle valve, air bypass, intake and exhaust valves.

The inertial phenomena can be considered as minor efforts but not complectly negligible. They describe the reduction or the increase of the pressure upstream the valve of a quantity proportional to the derivative of the mass flow through the same valve. Here, they are modeled as an linear inductance.The combustion process constitutes the most meaningful and complex phenomenon occurring into the engine. In order to model the in-cylinder cycle pressure, an equivalent electric circuit has been adopted. The circuit is formed by a variable condensator, representing the cylinder volume, equipped by an impulsive voltage generator. This causes an impulsive increase of the voltage at the condensator extremities and, consequently, generates a current flow thought the capacitor. This phenomenon corresponds to the well known combustion process, i.e. an impulsive increase of the in-cylinder pressure caused by the combustion resulting in a torque generation and in mass flow through the exhaust valves. Based on the electric analogies, the entire engine can be represented by the circuit shown in Figure 1.

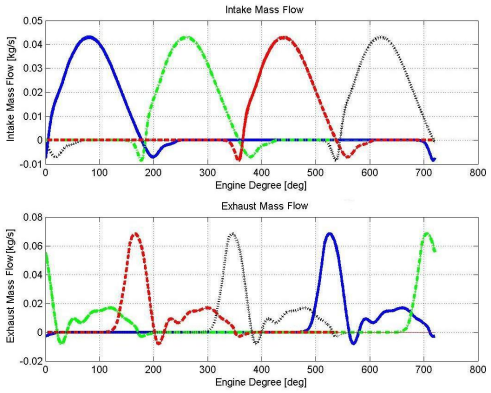


Fig. 2. Experiment 1 at 1500 rpm and WOT. Intake and exhaust mass flows for the four cylinders.

Starting from the left of the figure, the model describes the dynamic of the air crossing the intake manifold, i.e. driven by the ambient pressure (a current generator), the air mass passes the filter (a resistance) and the throttle body (a variable resistance) and arrives into the cylinders through the intake valves (a new variable resistance). The cylinders are then described by a parallel of ”n” combustion equivalent circuits, with ”n” the number of cylinders composing the engine.

In order to verify the reliability of the proposed model, some experiments are conducted aimed at comparing some measurable variables with experimental signals. The model has been designed by means of a simple approach based on the analogy of mechanical components with electrical circuits. This methodology has been presented in (7), where the validation of the model has been detailed.

Figures 2 and 3 report an experiment at 1500 rpm. In particular, 2 shows the simulated intake and exhaust mass flows for each cylinder. Figure 3 completes the experiment comparing the simulated in-cylinder pressure cycle with experimental data highlighting the good performance. The intake and exhaust valve lift ends the Figure 3.

The results, showing a good level of the model of reliability and accuracy, allows to carry on the simulated analysis of air mass dynamics for two different control strategies. It is remarked again that the purpose of the work is the study in simulation of phenomenon aimed at a better knowledge of the system.

III. Variable Valve Actuation System

The variable valve actuation has been introduced as a promising technology able to improve the performance of the vehicle in terms of fuel economy, emission reductions and, more generally, the whole efficiency of the system (4).

In opposition to the classical engine, where the intake and exhaust valves are commanded mechanically by the camshaft and so both the timing and the duration of valves opening are

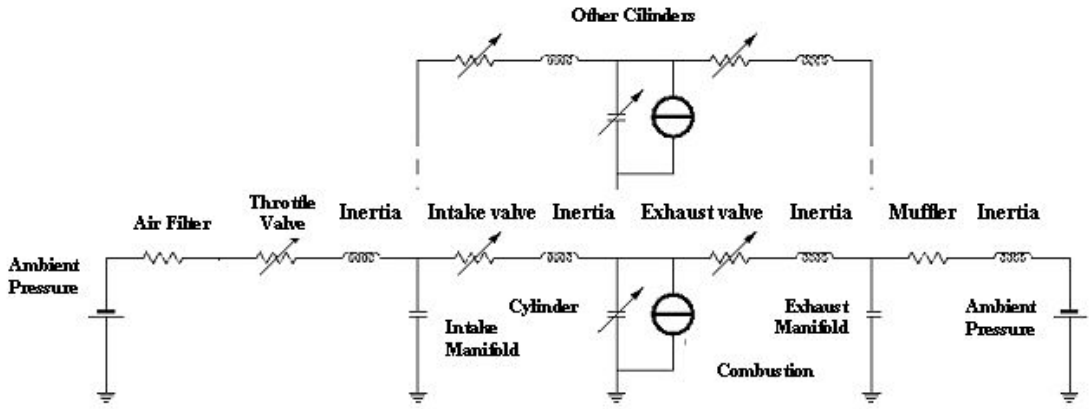


Fig. 1. Internal combustion engine equivalent circuit.

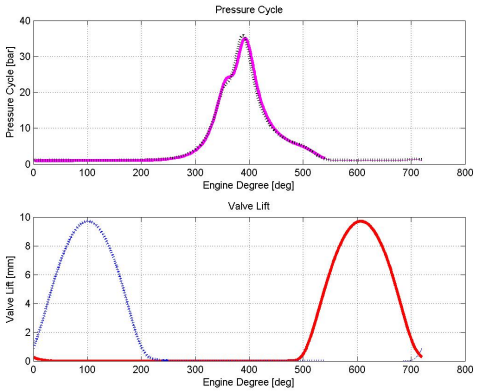


Fig. 3. Experiment 1 at 1500 rpm and WOT. The first plot compares experimental data (dotted-black line) of the pressure inside cylinder and simulated results (solid-magenta line); the second plot reproduces the intake and exhaust valves lift (experimental data).

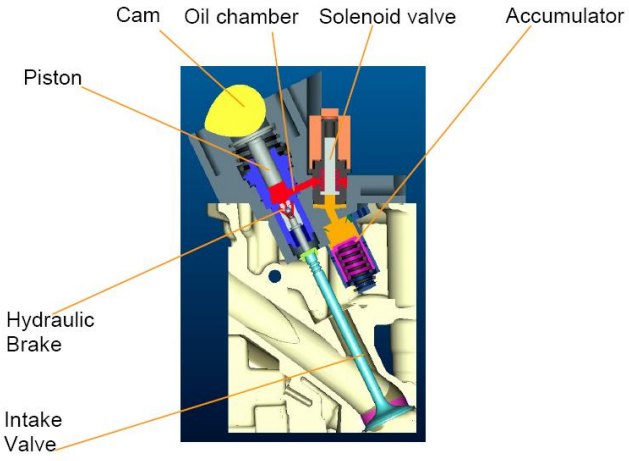


Fig. 4. VVA system.

fixed by events, the VVA system offers the possibility to vary the valves actuation.

The adopted VVA system is shown schematically in Figure 4. In this work, only the intake valves are actuated since the benefits to actuate as well the exhaust valves are considered small and does not justify the increase of costs to realize them. The valve actuator consists of a piston connected through an oil chamber to the intake valve, a solenoid valve to regulate the pressure inside the oil chamber and an hydraulic brake to assure the soft landing.

As reported in Figure 5, the system can operate in the following operating modes:

- Full Lift (FL) represents the normal functioning of the valves, i.e. commanded mechanically by the camshaft: the solenoid valve remains closed assuring high pressure into the oil chamber and, consequently, a rigid connection between the intake valve and the camshaft through the piston;
- Early Closure (EC) is obtained by opening the solenoid valve at a certain cam angle, i.e. the control angle, reducing the pressure inside the oil chamber. The motion of the intake valve is then decoupled from the piston and, forced by the valve springs, it starts to close earlier than in the full-lift mode. Soft landing of the intake valve is controlled by an hydraulic dampening unit (hydraulic brake);
- Late Opening (LO) can be achieved by regulating the solenoid valve partially opened. In this way, the pressure inside the oil chamber is regulated to a lower pressure than in the full lift mode, obtaining a rigid connection, but with a shorter distance function of the chamber pressure, between the intake valve and the camshaft. Consequently,

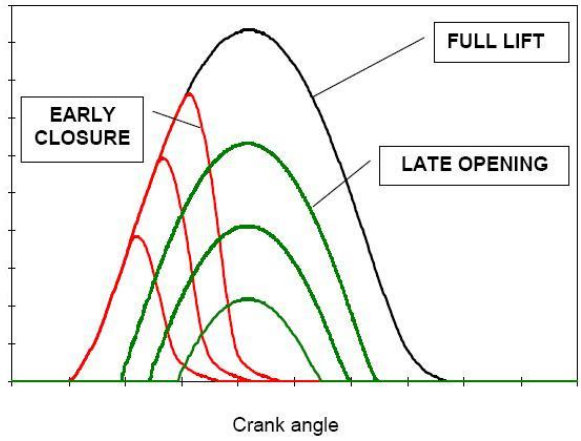


Fig. 5. Valve lift profiles. The valve lift values on the y-axes are omitted for confidential reason.

the valve profile is similar to the full lift mode, but with a smaller time duration;

- Multi Lift (ML) is a particular operative actuation mode obtained combining the late opening with the early closure. This profile is limited by the mechanical cam constrains, in fact the next late opening must be activate before of the 50% of the full lift cam.

The flexibility of intake valve control offered by the VVA system leads to enhance the efficiency of the combustion process. As an example, it is possible to deactivate one of the four valves per cylinder in order to produce swirls and so improve the combustion optimizing the propagation of the flame inside the cylinder. Or it is possible to actuate the early closure mode to reduce pumping losses and improve fuel economy. Or, again, to deactivate half cylinders to improve fuel economy, simultaneously increasing low-speed torque. More in general, the following advantages can be addressed to the introduction in the vehicle of the VVA system:

- high charge trapping efficiency over the entire speed range through a wide modulation of valve lift;
- throttle-less engine operation, through direct air control at the valves resulting in a reduction of pumping work and fuel consumption;
- dynamic control, cylinder by cylinder and stroke by stroke, of the inlet charge aimed at an improvement of emissions, driveability and fuel consumption in transient operation.

In this work, it is presented an analysis on the interaction between the VVA application and the idle control through the mid-ranging strategy. Idling control is one of the most important closed loop control functions for an internal combustion engine. In particular, modern engine control software provides the opportunity to stop the engine and cleverly to restart it when it is necessary, this function is called Stop and Start. This operation doesn't reduce the idle control relevance but, increases it, especially with the introduction of new devices such as CVCP (Continuously Variable Cam Phaser), VVA, Turbo groups and so on.

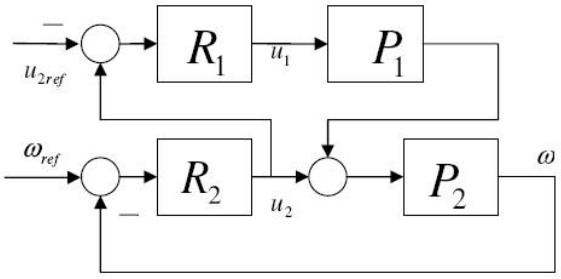


Fig. 6. Mid-ranging control scheme. A slow control loop driven by R1 forces the control input u2 to a steady-state desired value.

The presence of the VVA can cause uncertainty for engine fueling rate in order to control the air-fuel ratio during the combustion and, consequently, to optimize the performance of idle management. In fact the different profiles of the VVA, full lift, early closing, late opening and multi lift, determinate a variable air mass flow depending by VVA control, as described in details in the following. So it is critical to manage the engine speed and its fluctuations around the desired value and at designing suitable control strategies to compensate this uncertainty.

For the simulations here presented, the VVA on the engine has been conceived to operate only on the intake valves and in un-throttled mode, except for cut-off condition. About the VVA control mode here considered, it has been applied the Early Closing of the intake valve as shown in Figure 5. In this way, the intake valve closing reduces the fresh air charge for every single stroke. So, the charge, torque and power of the engine can be controlled only with VVA control.

IV. Mid-Ranging Idle Control

In the control field, there are numerous practical examples of control algorithms where, in order to meet the control objectives, two input must be manipulated to control a single output (9). In some cases, this may be achieved by manipulating one input at the time. Such strategies are often referred to as split ranging. In other situation, it may be desirable or even necessary to simultaneously manipulate the inputs. For example, consider the situation shown in Figure 6, where the speed ω is controlled by a combination of two controllers in parallel. In particular, a slow control loop driven by $R1$ forces the control input $u2$ to a steady-state desired value. This is called Mid-Ranging technique. Today, the Mid-Ranging technique is largely used for idle speed control of spark ignition engines, usually controlled by commercial Electronic Control Unit (ECU) based on "Torque Based" architecture (10).

Idling is one of the most often used functionalities in the modern car. This is especially the case in city traffic, where there are frequent stop and go situations. Therefore, improvements of the control performance for the idle speed control unit has always been a high priority. That is, keep

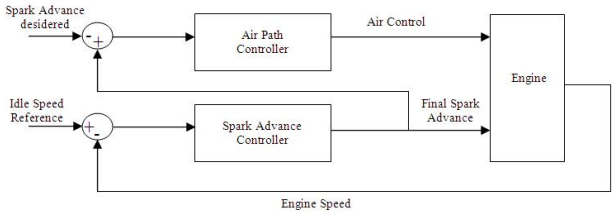


Fig. 7. The Mid Ranging Idle Speed Control scheme.

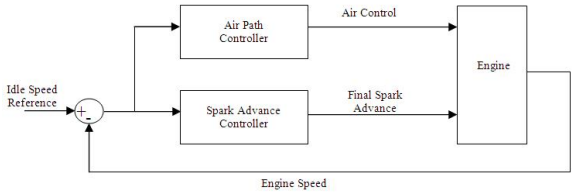


Fig. 8. The Modified Mid Ranging Idle Control Scheme.

the engine speed at a desired setpoint value, ensure good disturbance rejection while maintaining low fuel consumption. Typical disturbances that are to be rejected by the controller are loads from the air-conditioning system or power-steering. Obviously, the ECU compensates such disturbances on engine torque by using the throttle, however due to the slow dynamics of the air mass in the intake manifold, this would generate an unacceptably slow disturbance rejection. For this reason the spark advance is used as a second control signal, by advancing or retarding the ignition and obtaining an instantaneous torque variation from the engine. However, a deviation from the optimal spark ignition will result in higher fuel consumption. Thus the use of this signal should be kept at minimum and used only for improving the speed of the disturbance rejection. From control point of view this is a difficult problem since the system in question is nonlinear, multivariable (two inputs) and time varying. Moreover, the throttle control channel has a slower dynamics than the spark advance event. In the literature this problem is usually approached by treating the two control channels separately (one control signal is set to constant while the other is modified), leading to performance degradation. Some other approaches treat linearized models resulting in local designs. There exist approaches where both control signals are treated in the same time (multivariable control), however the resulting controllers are highly complex and difficult to tune. This article proposes the usage of a simple technique originally used in process control, called Mid-Ranging 'series', to analyze the idle speed control problem. The particular of this scheme is that the throttle is governed directly by the error between the desired and actual spark advance, while the spark advance is governed by the engine speed error, as reported in Figure 7. Moreover, the technique is particularly suitable for processes where one of the inputs has faster dynamics than the other, this is precisely true for the idle speed control problem in traditional SI engine. In the Figure 7 the spark advance control loop is the fast control loop, it takes as reference value the desired engine speed. The second loop contains the slower dynamics, where the air path controller adjusts the throttle angle such that in stationarity the spark advance will converge to the desired value. Traditionally the mid-ranging schemes are based on PID controllers, which will be used here too. This idea is correctly based on experimental observation, the time delay between the spark advance application and the engine torque response is relatively small, symmetrical and predictable, if compared with the engine torque response from the intake manifold filling/emptying gas dynamics. This control concept, depicted in Figure 7, is simple and robust for a standard spark ignition engine, as demonstrated in (10) and in commercial ECU present on the market.

V. Mid-Ranging Algorithm with VVA System

In this work, the authors, starting from the comparison with Mid-Ranging 'series' scheme, have realized an analogue control scheme for an engine model equipped with VVA system. In the traditional engine, the air path is considered the slow via to regulate the engine torque, otherwise, in the modern engine equipped with VVA system the engine torque response, for any valve closing change, is able to exhibit a time delay comparable with the engine torque response to spark advance variation. For this reason, the Mid Ranging control scheme has been changed from its traditional form with two input in a 'parallel' scheme with only an input, as shown in Figure 8. The advantage of this scheme is to avoid severe interaction between the inner and outer control loops of the serial scheme. In this way, both the control loops are driven from the same rpm error without control interactions, improving the advantages about the authority, quickness and steady state performance.

For the experiments has been used the engine model previously described with VVA application, and the same standard controllers for spark advance and air path. In this experiment, the controllers are able to guarantee the same tuning easiness as the serial scheme, without any loss of performance.

The experiments have been realized applying the control strategy to the four stroke internal combustion engine model. In Figure 9 is reported a simulation for engine speed based on two Mid-Ranging control schemes, the results define the quality and robustness of the idle speed control for both techniques. The good quality of the idle control and engine behavior has been obtained by using only a Proportional-Derivative controller for the spark advance control loop and a Proportional-Integral controller for the engine air path control loop.

In Figure 10 the spark advance behavior of two kind of control scheme is depicted, showing the good preserved robustness, slightly improved by the opportunity to use a more high spark advance during the engine work and so to improve the fuel consumption. Moreover, it should be noted that the general improvement partly depends by the better cylinder

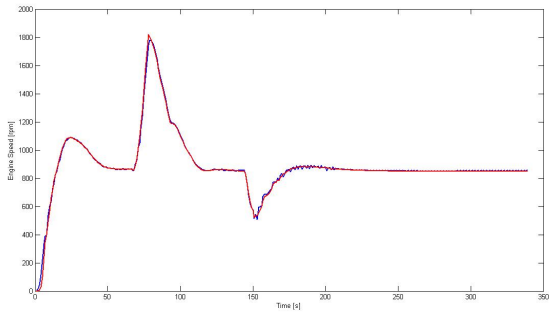


Fig. 9. Engine speed: comparison between the Mid-Ranging 'parallel' scheme (red-solid line) and Mid-Ranging 'series' scheme (blue-dotted line).

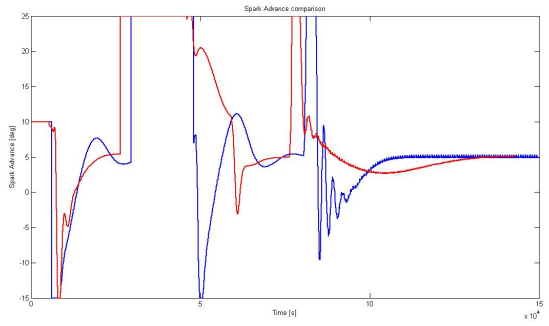


Fig. 10. Spark advance: comparison between the Mid-Ranging 'parallel' scheme (red-solid line) and Mid-Ranging 'series' scheme (blue-dotted line).

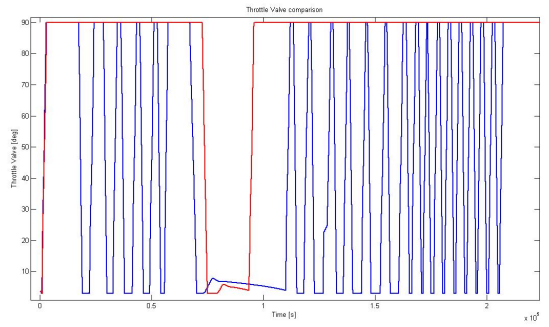


Fig. 11. Throttle valve opening: comparison between the Mid-Ranging 'parallel' scheme (red-solid line) and Mid-Ranging 'series' scheme (blue-dotted line).

filling efficiency, that un-throttled VVA systems usually allows, as reported in Figure 11, where is possible to note that the throttle valve is used only for cut-off occurrence.

VI. Conclusion

A standard Mid Ranging scheme for idle speed control in spark ignition engine has been investigated on a realistic four cylinder engine model. A similar control scheme, here named "parallel" Mid Ranging scheme, specific for engine equipped with VVA has been deducted and tested in similar conditions, showing similar and slightly better properties in terms of performance quality and tuning easiness. The ability of the parallel Mid Ranging control scheme for idle speed control in internal combustion engines inspire to think that even other kinds of control variables, such as instantaneous air/fuel ratio, could be added in order to obtain a more robust idle controller, without increases the tuning procedure. In such a way, the idle control for a spark ignited engine will be more hopefully able to fulfil the future requirements in terms of emission reduction, fuel consumption, disturb rejection, and driver satisfaction. With three charge/torque control variables (air, spark advance, air/fuel ratio) or even more (injection phase for direct fuel injection systems), the challenge will be the implementation of relatively simple, modular and separated control loop laws, with simple and almost separated tuning procedures.

References

[1] Z. Ye, *Modeling, identification, design, and implementation of nonlinear automotive idle speed control systemsan overview*, IEEE Transactions on Systems, MAN, and Cybernetics, Vol. 37, No. 6, 2007.

[2] C. Mariniello, M.T. Ragno, A. Palladino and G. Fiengo, *Fuel Consumption Reduction during Automotive Idle Speed Control*, European Control Conference, 2009.

[3] M. Thornhill, S. Thompson and H. Sindano, *A comparison of idle speed control schemes*, Control Engineering Practice, Vol. 8, No. 5, pp. 519-530, 2000.

[4] H.B. Lee, H. Kwon and K. Min, *Effects of various VVA systems on the engine fuel economy and optimization of a CVVT-VVL SI engine using 1D simulation*, International Journal of Automotive Technology, Vol. 8, No. 6, pp. 675-685, 2007.

[5] A. Palladino, G. Fiengo, F. De Cristofaro, A. Casavola and L. Glielmo, *In cylinder air charge prediction for VVA system: Experimental validation*, IEEE International Conference on Control Applications, No. 4629666, pp. 239-244, 2008.

[6] D. Pavkovic, J. Deur and I. Kolmanovsky, *Adaptive Kalman Filter-Based Load Torque Compensator for Improved SI Engine Idle Speed Control*, IEEE Transaction on Control System Technology, Vol. 17, No. 1, pp. 98-110, 2009.

[7] A. Palma, A. Palladino, G. Fiengo, F. De Cristofaro, F. Garofalo and L. Glielmo, *In cylinder air charge prediction for VVA system: Experimental validation* , IFAC World Conference, 2008.

[8] A. Zagrai, *Electro-mechanical analogies for modeling the structural impedance response*, Proceedings of SPIE - The International Society for Optical Engineering, No. 65320F, 2007.

[9] B. J. Allison and A.J. Isaksson, *Design and performance of midranging controllers*, Journal of Process Control, Volume 8, Issue 5, pp. 469474, 2009.

[10] S. Solyom and S. Eriksson, *Mid-Ranging Scheme for Idle Speed Control of SI Engines*, SAE Technical Paper, No. 2006-01-0608, 2006.

14 Optimal Control Design for Polynomial Nonlinear Systems using Sum of Squares Technique with Guaranteed Local Optimality

Optimal Control Design for Polynomial Nonlinear Systems using Sum of Squares Technique with Guaranteed Local Optimality

N. Boonnithivorakul and F. Pourboghrat, *Senior Member, IEEE*

***Abstract*—This paper deals with a computational approach to find the optimal control for nonlinear systems with polynomial vector fields. The approach involves four steps to find global optimality. First, local optimal control is found for the linearized part of the system and the quadratic part of the given performance index. Second, the density function method is used to find a stabilizing polynomial control for the nonlinear system. Third, the corresponding Lyapunov function is found for the control. Finally, the pair of control and its Lyapunov function are iteratively updated, using SOSTOOLS, for global optimal control. Numerical examples illustrate the effectiveness of the design approach.**

Keywords—nonlinear systems, optimal control, sum of squares

I. Introduction

In recent years, many researchers have shown interest in the sum of squares (SOS) technique introduced by Parrilo [1]. The fundamental method behind this technique is that the SOS problems can be converted into a convex optimization problem, which can be solved efficiently using semidefinite programming (SDP). This technique has been applied in many fields of control systems including stability analysis of a nonlinear systems [2] and [3]. Some control applications have also been discussed in [4] and [5].

In general, in order to obtain a nonlinear optimal control, one needs to solve the Hamilton Jacobi inequality corresponding to a given performance index [6]. Although, the exact solution may not be available, a local optimal control and its Lyapunov function exist if the nonlinear system has a controllable Jacobi linearization and the performance index has a Taylor series expansion with quadratic leading term [7] and [8].

Our objective is to obtain the nonlinear optimal control that matches the linear optimal control law for the linearized system to guarantee local optimality. The method in [9] can be used to solve the Hamilton Jacobi inequality simultaneously for finding both the nonlinear controller and its corresponding Lyapunov function, iteratively, where the nonlinear system is represented in a parameter dependent form. However, the original iterative algorithm requires an initial stabilizing controller and a corresponding Lyapunov function in order to be applied. The linear optimal control law and its Lyapunov function will not be iterated during the iterative algorithm.

Here, the density function method [10] is used to find an initial stabilizing polynomial controller, using SOS technique [2]. Some examples of this method have been shown in [11] and [12]. Next, using the Hamilton Jacobi inequality and the SOS method, a Lyapunov function is also found for the initial stabilizing control. Then, given a quadratic performance index (cost), an iterative procedure utilizing the SOS technique is considered so as to find the optimal stabilizing polynomial control that minimizes the cost.

Polynomial parameters used in this paper can be mainly separated into two parts which are linear and nonlinear parts. The linear part of the system is denoted by a subscript "*l*", such as $u_l(x)$, which is the optimal control for the linearized system. All higher order functions in the system are denoted with a subscript "*h*" such as $u_h(x)$, which is the nonlinear part of the optimal control.

The remainder of the paper is organized as follows. A brief introduction to the SOS formulation is presented in Section II. Locally optimal control design and its Lyapunov function are presented in Section III. Nonlinear controller design using the density function is discussed in Section IV. In Section V, the search for the Lyapunov function is discussed and then, given a cost, an iterative algorithm for global optimal control design is explained later in this section. Section VI presents examples. Finally, the conclusions are summarized in Section VII.

II. Sum Of Squares Polynomials

The main computational method used in this paper is based on sum of squares (SOS) decomposition of multivariate polynomials [1]. The multivariate polynomial $p(x) = p(x_1,...,x_n)$ is a sum of squares if there exist polynomials $f_i(x), i = 1,...m$ such that

$$p(x) = \sum_{i=1}^{m} f_i^2(x) \tag{1}$$

The existence of sum of squares representation for $p(x)$ is sufficient condition for its non-negativity by the following proposition.

Proposition 1, **[**2]: Let $p(x)$ be a polynomial in $x \in R^n$ of degree $2d$. In addition, let $Z(x)$ be a column vector, which is a properly chosen vector of monomials in x with polynomials of degree no greater than d. Then $p(x)$ is a

N. Boonnithivorakul is with the Department of Electrical and Computer Engineering, Southern Illinois University at Carbondale (SIUC), Carbondale, IL 62901-6603 USA (e-mail: nattpong@siu.edu).

F. Pourboghrat is with the Department of Electrical and Computer Engineering, Southern Illinois University at Carbondale (SIUC), Carbondale, IL 62901-6603 USA (e-mail: pour@siu.edu).

sum of squares iff there exists a positive symmetric semidefinite matrix Q such that

$$p(x) = Z^T(x)QZ(x) \quad (2)$$

When $p(x)$ is a sum of squares, matrix $Q \geq 0$ can be found using semi-definite programming (SDP) technique.

III. Locally Optimal Control Design

A. Problem formulation

Consider the nonlinear system in the form of

$$\dot{x} = f(x) + g(x)u \quad (3)$$

where $x \in R^n$ is the state and $u(x) \in R^m$ is the control input. We assume that $f(x)$ and $g(x)$ are sufficiently smooth and that $f(0) = 0$ and $g(0) = B_l$ where $B_l \in R$.

The objective is to find the optimal control, $u(x)$, with respect to a performance index

$$\vartheta(t) = \int_0^\infty (q(x) + Ru(x)^2)dt \quad (5)$$

In order to guarantee local optimality, we first separate the linear part of the system (3) using linearization method, which can be written as

$$\dot{x} = A_l x + f_h(x) + B_l(u_l + u_h) + g_h(x)(u_l + u_h) \quad (4)$$

where $A_l = \left.\frac{\partial f}{\partial x_i}\right|_{x=0}$, $f_h(x) = f(x) - A_l x$, $B_l = g(0)$

$g_h(x) = g(x) - B_l$ and $u(x) = u_l(x) + u_h(x)$.

We then decompose the integrand in (5) to include a quadratic term in x. Then (5) can be rewritten as [7]

$$\vartheta(t) = \int_0^\infty (x^T Q_l x + q_h(x) + R(u_l(x) + u_h(x))^2)dt \quad (6)$$

where $q(x)$ is positive-definite with

$$q_{xx}(0) = \frac{1}{2}\left.\frac{\partial^2 q(x)}{\partial x^2}\right|_{x=0} = Q_l, \text{ and } R > 0$$

B. Finding Local Optimal Control

The objective here is to find the locally optimal control u_l for the linear part of the system (3)

$$\dot{x} = A_l x + B_l u_l \quad (7)$$

to minimize the quadratic part of the performance index/cost

$$\vartheta_l(t) = \int_0^\infty (x^T Q_l x + Ru_l(x)^2)dt \quad (8)$$

We can then write the Hamilton Jacobi-Bellman (HJB) equation as

$$\frac{\partial V_l^T}{\partial x}(A_l x + Bu_l) + x^T Q_l x + Ru_l^2 = 0 \quad (9)$$

The optimal cost (the Lyapunov function) is given by [13]

$$V_l(x) = x^T P_l x \quad (10)$$

where, P_l is a symmetric and positive definite matrix. Thus, $\frac{\partial V}{\partial x} = 2P_l x$. Moreover, (9) can be reduced to the standard Riccati equation as

$$x^T(A_l^T P_l + P_l A_l + Q_l - P_l B_l R^{-1} B_l^T P_l)x = 0 \quad (11)$$

which is true for all x and reduces to the Algebraic Riccati Equation (ARE)

$$A_l^T P_l + P_l A_l + Q_l - P_l B_l R^{-1} B_l^T P_l = 0 \quad (12)$$

Finally, the local optimal control will be given by

$$u_l(x) = K_l x \quad (13)$$

where, $K_l = -R^{-1} B_l^T P_l$

IV. Stabilizing Nonlinear Control Design

A nonlinear control design used in this paper is based on density function, a dual of Lyapunov's stability theorem, which was proposed by Rantzer [10]. In general, for a nonlinear system, a major difficulty is a lack of a symmetric method in finding a stabilizing control $u(x)$ and a corresponding Lyapunov function $V(x)$ simultaneously. The problem is that, for a nonlinear system $\dot{x} = f(x) + g(x)u(x)$, the set $\{V, u\}$ satisfying the Lyapunov inequality

$$\nabla V[f(x) + g(x)u(x)] \quad < \quad 0 \quad (14)$$

is generally not convex, where $\nabla V = \frac{\partial V}{\partial x} = [\frac{\partial V}{\partial x_i}, ..., \frac{\partial V}{\partial x_n}]^T$.

However, in [10] it is shown that a dual problem, based on density function ρ, has much better convexity properties. In fact, the set $\{\rho, u\rho\}$ satisfying the inequality

$$\nabla \cdot [\rho(x)(f(x) + g(x)u(x))] \; > \; 0 \quad (15)$$

is convex, where by definition $\nabla \cdot f = \sum_{i=1}^n \frac{\partial f}{\partial x_i}$.

Hence, for systems with polynomial vector fields, we can use the SOS technique to find the pair $\{\rho, u\rho\}$ satisfying the inequality (15). The design strategy is based on the following result [10].

Theorem 1, [10]: Given the equation $\dot{x}(t) = f(x(t))$ where $f \in C^1(R^n, R^n)$ and $f(0) = 0$, suppose there exists a non-negative function $\rho \in C^1(R^n \setminus \{0\}, R^n)$ such that the term

$$\rho(x)f(x) / \|x\| \quad (16)$$

is integrable on $\{x \in R^n : \|x\| \geq 1\}$ and for almost all x

$$\nabla \cdot [\rho(x)f(x)] \quad > \quad 0 \quad (17)$$

Then, for almost all initial states $x(0)$ the trajectory $x(t)$ exists for $t \in [0, \infty)$ and tends to zero as $t \to \infty$. Moreover, if the equilibrium $x = 0$ is stable, then the conclusion remains valid even if ρ takes negative values. ■

To apply this theorem to find a nonlinear stabilizing control, let us first consider the system (4) with $u_l(x) = -R^{-1} B_l^T P_l x$, then we have

$$\dot{x} = A_l x + f_h - B_l R^{-1} B_l^T P_l x - g_h R^{-1} B_l^T P_l x + B_l u_h + g_h u_h \quad (18)$$

If we let

$$F(x) = A_l x + f_h - B_l R^{-1} B_l^T P_l x - g_h R^{-1} B_l^T P_l x \quad (19)$$

then, since $g(x) = B_l + g_h(x)$, (18) can be rewritten as

$$\dot{x} = F(x) + g(x)u_h \tag{20}$$

Consider the following parameterization for $\{\rho, u_h\rho\}$ [11]:

$$\rho(x) = \frac{a(x)}{b(x)^\alpha}, \quad u_h(x)\rho(x) = \frac{c(x)}{b(x)^\alpha} \tag{21}$$

here $a(x)$ and $b(x)$ are positive polynomials, $c(x)$ is a polynomial and α is chosen such that the integreability condition (16) in Theorem 1 is satisfied. Then (15) can be written as

$$\begin{aligned}\nabla\cdot[\rho(F+gu_h)] &= \nabla\cdot[\frac{1}{b^\alpha}(Fa+gc)] \\ &= \frac{1}{b^{\alpha+1}}[b\nabla\cdot(Fa+gc) - \alpha\nabla b\cdot(Fa+gc)]\end{aligned} \tag{22}$$

Since $b(x)$ is positive, we only need to satisfy

$$b\nabla\cdot(Fa+gc) - \alpha\nabla b\cdot(Fa+gc) \quad > \quad 0 \tag{23}$$

Therefore, (23) is in polynomial form and we can use SDP to search for the solution that makes the left-hand side a sum of squares. A stabilizing control can then be defined as

$$u_h(x) = \tfrac{c(x)}{a(x)} \tag{24}$$

Moreover, if $a(x)$ is chosen to be a constant real number, (e.g., $\alpha = 1$), then from (21), the corresponding stabilizing control will be a polynomial, given as

$$u_h(x) = c(x) \tag{25}$$

V. Optimal Polynomial Control Design

A. Finding the Lyapunov Function

Let us define system (4) in the state dependent form as [2]

$$\dot{x} = A(x)Z(x) + B(x)u(x) \tag{26}$$

where $A(x)$ and $B(x)$ are polynomial matrices in x, and $Z(x)$ is an $N\times 1$ vector of monomials in x, (e.g., terms of the form $x_1^\alpha x_2^\alpha \cdots x_n^\alpha$), satisfying the following assumption.

Assumption 1: $Z(x) = 0$ iff $x = 0$.

Now, let $M(x)$ be a $N\times n$ polynomial matrix whose (i, j)-th entry is given by

$$M_{ij}(x) = \frac{\partial Z_i}{\partial x_j}(x) \tag{27}$$

where $i = 1, ..., N$, $j = 1, ..., n$, and also denote

$$J = \{j_1, j_2, ..., j_m \mid j^{th} \text{ row of } B(x) \text{ equals } 0\} \tag{28}$$

Our primary objective here is to find a Lyapunov function corresponding to the nonlinear terms, $V_h(x)$, that satisfies the Hamilton Jacobi inequality for the control

$$u(x) = u_l(x) + u_h(x) \tag{29}$$

However, note that (29) can also be written in the form of

$$u(x) = K(x)Z(x) \tag{30}$$

Hence, the closed-loop system can be written as

$$\dot{x} = [A(x) + B(x)K(x)]Z(x) \tag{31}$$

Moreover, the performance index can also be written as

$$\vartheta(t) = \int_0^\infty (Z(x)^T QZ(x) + u^T Ru)dt \tag{32}$$

where Q and R are symmetric positive definite matrices.

Consider a Lyapunov function that consists of two terms

$$V(x) = V_l(x) + V_h(x) \quad > \quad 0 \tag{33}$$

where $V_l(x)$ found in Section III and $V_h(x)$ represents the Lyapunov function for the nonlinear part.

Therefore, (33) can be rewritten in the form of a parameter-dependent Lyapunov function as

$$V(x) = Z^T(x)P(\tilde{x})Z(x) \quad > \quad 0 \tag{34}$$

where $P(\tilde{x})$ is a symmetric positive definite polynomial matrix and $\tilde{x} = (x_{j_1}, x_{j_2}, ..., x_{j_m})$ with $j_1, j_2, ..., j_m \in J$ as stated in (28). Finally, let $A_j(x)$ denote the j-th row of $A(x)$.

Therefore, we have

$$\dot{V}(x) = \dot{Z}^T PZ + Z^T \dot{P}Z + Z^T P\dot{Z} \tag{35}$$

Then for the function $V(x)$ in (35) to represent the minimum of the performance index (32) over the set of all control functions $u(x)$, the expression for $\dot{V}(x)$ in (35) should be written as

$$\begin{aligned}\frac{\partial V}{\partial t}(x(t)) = Z^T[&(A+BK)^T M^T P + PM(A+BK) \\ &+ Q + K^T RK + \sum_{j=1}^m \frac{\partial P}{\partial x_j}(A_j Z)]Z \quad < \quad 0\end{aligned} \tag{36}$$

or equivalently as,

$$(A+BK)^T M^T P + PM(A+BK) + Q + K^T RK + \sum_{j=1}^m \frac{\partial P}{\partial x_j}(A_j Z) \quad < \quad 0 \tag{37}$$

If the inequalities (34) and (37) are satisfied, the closed-loop system (31) will be asymptotically stable at zero equilibrium. The polynomial inequalities (34) and (37) can now be viewed as sum of squares constraints, which can be solved for P and K using SOSTOOLS [14].

B. Pre-procedure for the Iterative Algorithm

Let us define system (23) in a parameterized form [9], as

$$\dot{x} = A(x,\theta)Z(x) + B(x,\theta)\,u \tag{38}$$

where $A(x,\theta)$ and $B(x,\theta)$ are polynomial matrices in x, as

$$\begin{aligned}A(x,\theta) &= A_1(x)\theta_1 + A_2(x)\theta_2 \\ B(x,\theta) &= B_1(x)\theta_1 + B_2(x)\theta_2\end{aligned} \tag{39}$$

such that the uncertain parameters $\theta = \begin{bmatrix}\theta_1 & \theta_2\end{bmatrix}^T \in R^2$ are constant and satisfy $\theta_1 \geq 0, \theta_2 \geq 0$ and $\theta_1 + \theta_2 = 1$. The corresponding Lyapunov function also is rewritten as

$$V(x,\theta) = Z(x)^T P(\tilde{x},\theta)Z(x) \tag{40}$$

where $P(\tilde{x},\theta) = P_1(\tilde{x})\theta_1 + P_2(\tilde{x})\theta_2$. Moreover, let

$$\Phi(x,\theta) = (A(x,\theta) + B(x,\theta)K(x))^T M(x)^T P(\tilde{x},\theta) + P(\tilde{x},\theta)M(x)(A(x,\theta) + B(x,\theta)K(x,\theta)) + \sum_{i=1}^{2}\sum_{j=1}^{m}\frac{\partial P_i}{\partial x_j}(\tilde{x},\theta)(A_{ij}(x,\theta)Z(x)) + Q \quad (41)$$

Also apply Schur complement to (37), then we get

$$\begin{bmatrix} \Phi(x,\theta) & K^T(x) \\ K(x) & -R^{-1} \end{bmatrix} < 0 \quad (42)$$

Let us define

$$N_i = \tilde{A}_i^T\tilde{P}_i + \tilde{P}_i\tilde{A}_i + \nabla\tilde{P}_i + \tilde{Q} \quad (43)$$

for i = 1,2, and

$$N_3 = \tilde{A}_1^T\tilde{P}_2 + \tilde{P}_2\tilde{A}_1 + \tilde{A}_2^T\tilde{P}_1 + \tilde{P}_1\tilde{A}_2 + \nabla\tilde{P}_3 + 2\tilde{Q} \quad (44)$$

where $\tilde{A}_i = \begin{bmatrix} M[A_i + B_iK] & 0 \\ 0 & 0 \end{bmatrix}$, $\tilde{P}_i = \begin{bmatrix} P_i & 0 \\ 0 & 0 \end{bmatrix}$, $\tilde{Q} = \begin{bmatrix} Q & K^T \\ K & -R^{-1} \end{bmatrix}$,

$\nabla\tilde{P}_i = \begin{bmatrix} \sum_{j=1}^{m}\left\{\frac{\partial P_i}{\partial x_j}(A_{ij}Z)\right\} & 0 \\ 0 & 0 \end{bmatrix}$, $\nabla\tilde{P}_3 = \begin{bmatrix} \sum_{j=1}^{m}\left\{\frac{\partial P_1}{\partial x_j}(A_{2j}Z) + \frac{\partial P_2}{\partial x_j}(A_{1j}Z)\right\} & 0 \\ 0 & 0 \end{bmatrix}$,

for i = 1,2. Then (42) can be written as

$$\theta_1^2 N_1 + \theta_2^2 N_2 + \theta_1\theta_2 N_3 < 0 \quad (45)$$

However, (45) is clearly not convex. Therefore, let us transform (45) into two separate inequalities as stated in Theorem 2 given below.

Theorem 2, [9]: Assume that there exist a polynomial control $u(x) = K(x)Z(x)$ and symmetric polynomial matrices $P_1(\tilde{x}) > 0$, $P_2(\tilde{x}) > 0$ and $Y_s(x) > 0$ such that

$$N_3 - Y_s < 0 \quad (46)$$

$$\begin{bmatrix} N_1 & Y_s/2 \\ Y_s/2 & N_2 \end{bmatrix} < 0 \quad (47)$$

where N_1, N_2, N_3 are given in (43) and (44). Then the closed-loop system (31) is asymptotically stable.

Proof: see [9]

However, the matrix inequalities (46) and (47) are still not convex because they still include the product of $K(x)$ and $P(\tilde{x},\theta)$. Theorem 3 provides a relaxed method of solving this problem by adding additional positive semi-definite polynomial matrices which tend to zero at the end of the iterative algorithm. Let us define the following positive semi-definite polynomial matrices.

$$\Omega_1 = 2\Delta_1\Delta_1^T + \Delta_2 B_2 B_2^T \Delta_2^T + \Delta_3 B_1 B_1^T \Delta_3^T \quad (48)$$

$$\Omega_2 = \begin{bmatrix} \Delta_1 \\ 0 \end{bmatrix}\begin{bmatrix} \Delta_1 \\ 0 \end{bmatrix}^T + \begin{bmatrix} 0 \\ \Delta_1 \end{bmatrix}\begin{bmatrix} 0 \\ \Delta_1 \end{bmatrix}^T + \begin{bmatrix} \Delta_2 \\ 0 \end{bmatrix} B_1 B_1^T \begin{bmatrix} \Delta_2 \\ 0 \end{bmatrix}^T + \begin{bmatrix} 0 \\ \Delta_3 \end{bmatrix} B_2 B_2^T \begin{bmatrix} 0 \\ \Delta_3 \end{bmatrix}^T \quad (49)$$

where $\Delta_1 = \begin{bmatrix} (K - K_0) & 0 \end{bmatrix}^T$, $\Delta_2 = \begin{bmatrix} (P_1 - P_{10}) & 0 \end{bmatrix}^T M$, and $\Delta_3 = \begin{bmatrix} (P_2 - P_{20}) & 0 \end{bmatrix}^T M$. By adding the above positive semi-definite polynomial matrices (48) and (49) to inequalities (46) and (47), respectively, we get

$$N_3 - Y + \Omega_1 < 0 \quad (50)$$

$$\begin{bmatrix} N_1 & Y/2 \\ Y/2 & N_2 \end{bmatrix} + \Omega_2 < 0 \quad (51)$$

where

$$Y = \begin{bmatrix} Y_{11}(x) & Y_{21}^T(x) \\ Y_{21}(x) & Y_{22}(x) \end{bmatrix} \quad (52)$$

It is clear that if (50) and (51) are satisfied, then (46) and (47) are satisfied. Let us define the following symmetric matrices

$$\Sigma_1 = \begin{bmatrix} E_{11} - Y_{11} & * & * & * \\ 2K - Y_{21} & -2R^{-1} - Y_{22} & * & * \\ K + B_1^T M^T P_2 & 0 & -I & * \\ K + B_2^T M^T P_1 & 0 & 0 & -I \end{bmatrix} \quad (53)$$

and

$$\Sigma_2 = \begin{bmatrix} E_1 & * & * & * & * & * \\ K & -R^{-1} & * & * & * & * \\ Y_{11}/2 & Y_{21}^T/2 & E_2 & * & * & * \\ Y_{21}/2 & Y_{22}/2 & K & -R^{-1} & * & * \\ K + B_1^T M^T P_1 & 0 & 0 & 0 & -I & * \\ 0 & 0 & K + B_2^T M^T P_2 & 0 & 0 & -I \end{bmatrix} \quad (54)$$

where * indicates the symmetric entries of the matrices, and

$$\begin{aligned} E_{11} &= A_1^T M^T P_2 + P_2 M A_1 + A_2^T M^T P_1 + P_1 M A_2 \\ &+ \sum_{j=1}^{m}\left\{\frac{\partial P_1}{\partial x_j}(A_{2j}Z)\right\} + \sum_{j=1}^{m}\left\{\frac{\partial P_2}{\partial x_j}(A_{1j}Z)\right\} + 2Q \\ &- P_1 M B_2 B_2^T M^T P_{10} - P_{10} M B_2 B_2^T M^T P_1 - 2K^T K_0 \\ &- P_2 M B_1 B_1^T M^T P_{20} - P_{20} M B_1 B_1^T M^T P_2 - 2K_0^T K \\ &+ P_{10} M B_2 B_2^T M^T P_{10} + P_{20} M B_1 B_1^T M^T P_{20} + 2K_0^T K_0 \end{aligned} \quad (55)$$

and

$$\begin{aligned} E_i &= A_i^T M^T P_i + P_i M A_i + \sum_{j=1}^{m}\left\{\frac{\partial P_i}{\partial x_j}(A_{ij}Z)\right\} \\ &- P_i M B_i B_i^T M^T P_{i0} - P_{i0} M B_i B_i^T M^T P_i \\ &- K^T K_0 - K_0^T K + P_{i0} M B_i B_i^T M^T P_{i0} + K_0^T K_0 + Q \end{aligned} \quad (56)$$

for i = 1,2. Now we have the following theorem.

Theorem 3, [9]: Consider the nonlinear system (38) and the performance index/cost (32). Given a stabilizing polynomial control gain matrix $K_0(x)$ and a corresponding Lyapunov function with symmetric positive definite matrices $P_{10}(\tilde{x})$ and $P_{20}(\tilde{x})$, suppose that there exist a different set of polynomial control gain matrix $K(x)$ and symmetric polynomial matrices $P_1(\tilde{x}) > 0$, $P_2(\tilde{x}) > 0$, and $Y(x) > 0$ such that

$$\Sigma_1(P_1, P_2, K, Y, P_{10}, P_{20}, K_0) < 0 \quad (57)$$

$$\Sigma_2(P_1, P_2, K, Y, P_{10}, P_{20}, K_0) < 0 \quad (58)$$

Then the closed-loop system (31) is asymptotically stable and the corresponding cost (32) will be smaller for all x.

Proof: see [9]

By applying Schur complement to the inequalities (50) and

(51), it is clear that the inequalities (57) and (58) are convex and are equivalent to (50) and (51), respectively. Therefore we can now use semidefinite programming to find the solution for $K(x)$ and $V(x)$, simultaneously.

C. Iterative algorithm procedure

Step 1: Find the linearized part of system (3), and find the local optimal control $u_l(x)$ corresponding to the quadratic part of the performance index/cost (8) in Section III

Step 2: Define the degree and form of the polynomial for the nonlinear controller $u_h(x)$, then find $u_h(x)$ using the density function shown in Section IV.

Step 3: Define the degree of the polynomial for $V_h(x)$, and consider $V(x) = V_l(x) + V_h(x) = Z^T(x)P(\tilde{x})Z(x)$ for properly selected basis $Z(x)$. Also consider $u(x) = u_l(x) + u_h(x) = K(x)Z(x)$. Then find $P(\tilde{x})$, and equivalently $V_h(x)$, using Lyapunov stability analysis stated in Section V.A.

Step 4: Let $P_{10}^1(\tilde{x}) = P_{20}^1(\tilde{x}) = P(\tilde{x})$ and $K_0^1(x) = K(x)$, then at the *i*-th iteration, minimize [*trace*($P_1^i(\tilde{x})$)] subject to $P_1(\tilde{x}) > 0, P_2(\tilde{x}) > 0$, the SOS conditions (52), (57), and (58). Note that, the parameters iterated in this step are only parameters corresponding to $u_h(x)$ and $V_h(x)$ which are embedded in $u(x)$ and $V(x)$, respectively. Parameters corresponding to $u_l(x)$ and $V_l(x)$ will be fixed at all time during the iteration. Then, stop the iteration if | *trace*($P_{10}^i(\tilde{x})$) - *trace*($P_{10}^{i-1}(\tilde{x})$) | < ε, where ε is some small positive real number.

VI. Simulation Results

A. Example 1

Consider a 2nd-order polynomial nonlinear system [15].

$$\begin{aligned} \dot{x}_1 &= x_2 \\ \dot{x}_2 &= x_1^3 + u \end{aligned} \tag{59}$$

Also, consider a performance index (5) with $q(x) = \frac{1}{2}(x_1^2 + x_2^2)$ and $R = 0.5$. Note that for this system the optimal control is known exactly in analytical form [15]. However, the goal is to find the optimal control using the proposed method.

Step 1: Find the optimal control and its Lyapunov function for the linearized system, as

$$u_l = -x_1 - 1.732x_2,\ V_l = 0.866x_1^2 + x_1x_2 + 0.866x_2^2 \tag{60}$$

Step 2: Define the stabilizing polynomial control to be of 3rd-order as $u_h = k_1x_1^3 + k_2x_1^2x_2 + k_3x_1x_2^2 + k_4x_2^3$. Let $a(x) = 1$, $\alpha = 8$, and $b(x) = x_1 + x_2$ for the density function. Then solve the inequality (23) for k_j, $j = 1,...,4$, using SOSTOOLS, to obtain

$$k_j = [-5.936,\ 4.270,\ 4.270,\ -4.936] \tag{61}$$

Step 3: Choose $V_h(x)$ as a polynomial of 4th-order and select $Z(x) = \begin{bmatrix} x_1 & x_2 \end{bmatrix}^T$, so that $V(x) = V_l + V_h = Z(x)^T P(\tilde{x})Z(x)$, where

$$P = \begin{bmatrix} 0.866 + p_1x_1^2 + p_2x_1x_2 + p_3x_2^2 & 0.5 + p_4x_1^2 + p_5x_1x_2 + p_6x_2^2 \\ 0.5 + p_4x_1^2 + p_5x_1x_2 + p_6x_2^2 & 0.866 + p_7x_1^2 + p_8x_1x_2 + p_9x_2^2 \end{bmatrix}$$

and $p_i \in R, i = 1,...,9$. Also let $u(x) = u_l + u_h = K(x)Z(x)$. Then the parameters p_i, $i = 1,...,9$, in $P(\tilde{x}) = P(x, V_l, p_i)$ can be found, using SOSTOOLS, to satisfy the inequality (37), which results in

$$\begin{aligned} p_i = [&4.484,\ -2.904,\ 0.828,\ 2.322,\ 3.062, \\ &-6.636,\ 2.235,\ 9.608,\ 2.637] \end{aligned} \tag{62}$$

Step 4: Select $\theta_1 = \theta_2 = 0.5$ and let $P_{10}^1(\tilde{x}) = P_{20}^1(\tilde{x}) = P(\tilde{x})$ and $K_0^1(x) = K(x)$. The optimal control can be found by finding new sets of matrices $P_1(\tilde{x})$, $P_2(\tilde{x})$, and $K(x)$, iteratively, using SOSTOOLS, that satisfy the inequalities

$$\sum_1 > 0, \sum_2 > 0, P_1(\tilde{x}) > 0, P_2(\tilde{x}) > 0, Y(x) > 0 \tag{63}$$

After 10 iterations, the final results for $K(x)$ and $P(\tilde{x}) = P(x, V_l, p_i)$ for $, i = 1,...,9$ are found as

$$K(x)^T = \begin{bmatrix} -1 - 0.531x_1^2 - 0.116x_1x_2 - 0.014x_2^2 \\ -1.732 - 0.116x_1^2 - 0.014x_1x_2 + 0.126x_2^2 \end{bmatrix} \tag{64}$$

and

$$\begin{aligned} p_i = [&3.109,\ -3.066,\ 1.231,\ 1.507,\ 1.192, \\ &-0.496,\ -0.040,\ 0.540,\ 1.630] \end{aligned} \tag{65}$$

with the optimal polynomial control given by

$$\begin{aligned} u(x) = &-x_1 - 1.732x_2 - 0.531x_1^3 - 0.232x_1^2x_2 \\ &- 0.027x_1x_2^2 - 0.126x_2^3 \end{aligned} \tag{66}$$

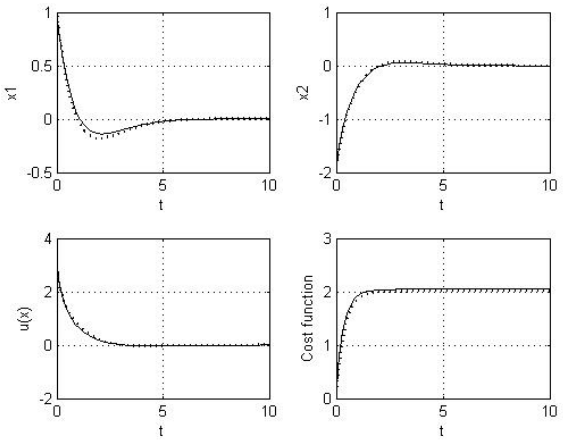


Figure 1. Comparison of the control, states, and cost between the exact solution (dashed line) and the proposed solution (solid line)

For the initial states $x(0) = [1, -2]$, Figure 1 shows the optimal control, the states, and the optimal cost for the proposed algorithm as well as those corresponding to the

exact optimal control, which is given by [15]

$$u^*(x) = -x_1^3 - x_1\sqrt{x_1^4+1} - x_2\sqrt{2x_1^2 + 2\sqrt{x_1^4+1}+1} \qquad (67)$$

Clearly the optimal solution using the proposed method is very close to the exact optimal control solution (67).

B. Example 2

Consider a polynomial nonlinear system with dynamics

$$\begin{aligned} \dot{x}_1 &= x_2 - x_1^3 + x_1^2 + 0.5x_1^2x_2 + 2x_2^3 \\ \dot{x}_2 &= x_1 + u \end{aligned} \qquad (68)$$

The optimal control is to be found so as to minimize the performance index

$$\vartheta(t) = \int_0^\infty (x_1^2 + x_2^2 + u^2)dt \qquad (69)$$

We first linearize system (68) and find the optimal control and the Lyapunov. Next, using density function technique with $a(x) = 1$, $\alpha = 10$, and $b(x) = (x_1 + x_2)$, find the control in the form of $u_h = k_1x_1^3 + k_2x_1^2x_2 + k_3x_1x_2^2 + k_4x_2^3$, using SOSTOOLS. Then, form $u(x) = u_l + u_h = K(x)Z(x)$ and define the polynomial $V_h(x)$ to be of 4th-order and select $Z(x) = \begin{bmatrix} x_1 & x_2 \end{bmatrix}^T$. Then, use SOSTOOLS to find $P(\tilde{x})$ so that $V(x) = Z(x)^T P(\tilde{x})Z(x) > 0$. Next, let $\theta_1 = \theta_2 = 0.5$, $P_{10}^1(\tilde{x}) = P_{20}^1(\tilde{x}) = P(\tilde{x})$, and $K_0^1(x) = K(x)$. The optimal control is then determined by finding new sets of matrices $P_1(\tilde{x})$, $P_2(\tilde{x})$, and $K(x)$, iteratively, using SOSTOOLS, that would satisfy the inequality constraints (63). After 14 iterations, the final control is

$$\begin{aligned} u(x) = &-2.414x_1 - 2.414x_2 - 0.527x_1^3 - 0.235x_1^2x_2 \\ &+ 0.157x_1x_2^2 + 0.211x_2^3 \end{aligned} \qquad (70)$$

Figure 2 shows the results of the proposed algorithm for $x(0) = [-3, 2]$.

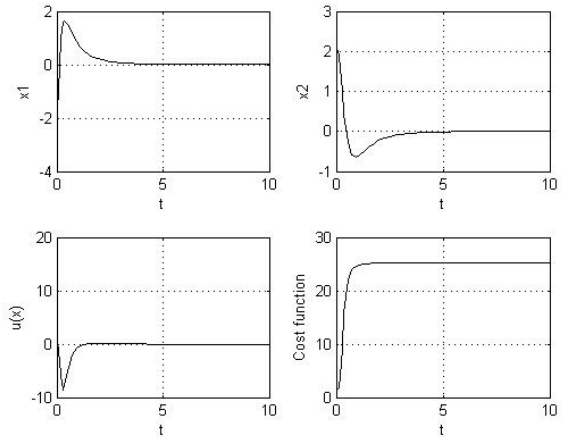


Figure 2. Solutions of the proposed algorithm

VII. Conclusions

In this paper, a computational procedure is developed to solve the nonlinear optimal control problem with guaranteed local optimality. The procedure begins with finding the local optimal control for the linearized system and the quadratic part of the given performance index. Then a nonlinear stabilizing polynomial control and its corresponding Lyapunov function are found using density functions. Subsequently, an iterative algorithm is employed to find a nonlinear optimal control that minimizes the given performance index (cost). Consequently, the proposed procedure results in an approximate global optimal control with guaranteed local optimality.

References

[1] P.A. Parrilo, "Structured Semidefinite Programs and Semialgebraic Geometry Methods in Robustness and Optimization," Ph.D. thesis, California Institute of Technology, Pasadena, CA, 2000.

[2] A. Papachristodoulou and S. Prajna, "A Tutorial on Sum of Squares Techniques for Systems Analysis", *American Control Conference*, Portland, OR, Jun. 8-10, 2005.

[3] S. Prajna, A. Papachristodoulou, and F. Wu, "Nonlinear Control Synthesis by Sum of Squares Optimization: A Lyapunov-based Approach," *Asian Control Conference,* vol. 1, pp. 157-165, Jul. 2004.

[4] Z. Jarvis-Wloszek, R. Feeley, W. Tan, K. Sun, and A. Packard, "Some Controls Applications of Sum of Squares Programming," *IEEE Conference on Decision and Control,* vol. 5, pp. 4676-4681, Maui, HI, Dec. 2003.

[5] M. de Oliveira, "Decomposition of a Polynomial as a Sum-of-Squares of Polynomials and the S-Procedure," *IEEE Conference on Decision and Control,* pp. 1654-1659, Seville, Spain, Dec. 12-15, 2005.

[6] A.J. Van der Schaft, "L -gain analysis of nonlinear systems and nonlinear state feedback H_∞ control," *IEEE Transactions on Automatic Control*, vol. 37, pp. 770-784, 1992.

[7] K. Ezal, Z. Pan, and P.V. Kokotović, "Locally Optimal and Robust Backstepping Design," *IEEE Transactions on Automatic Control,* vol. 45, no. 2, pp. 260-271, Feb. 2000.

[8] Z. Pan, K. Ezal, A.J.Krener, and P.V. Kokotović, "Backstepping Design with Local Optimality Matching," *Proceedings of the American Control Conference,* Arlington, VA, Jun. 25-27, 2001.

[9] D. Zhao, J.L. Wang, F. Liao, and E.K. Poh, "Nonlinear Optimal Control for Parameter-Dependent Polynomial Nonlinear Systems," *IEEE International Conference on Control and Automation*, pp. 1710-1715, Guangzhou, China, May 30-Jun. 1, 2007

[10] A. Rantzer, "A dual to Lyapunov's stability theorem", System and Control Letter, vol. 42, no. 3, pp166-168(8), Mar. 2000.

[11] S. Prajna, P.A. Parrilo, A. Rantzer, "Nonlinear Control Synthesis by Convex Optimization," *IEEE Transactions on Automatic Control,* vol. 49, no. 2, pp. 310-314, Feb. 2004.

[12] A. Ataei-Esfahani and Q. Wang, "Nonlinear Control Design of a Hypersonic Aircraft Using Sum-of-Squares Method," *American Control Conference*, New York, Jul. 11-13, 2007.

[13] C.T. Abdallah, V. Cerone, and P. Dorato, *Linear Quadratic Control: An Introduction*. Malabar, FL: Krieger Publishing Company, 2000.

[14] S. Prajna, A. Papachristodoulou, P. Seiler, and P.A. Parrilo, SOSTOOLS: Sum of Squares Optimization Toolbox for Matlab 2004 [Online]. Available: http://www.cds.caltech.edu/sostools/

[15] H.T. Banks, B.M. Lewis, H.T. Tan, "Nonlinear Feedback Controllers and Compensators: A State-Dependent Riccati Equation Approach," J. Computational Optimization and Applications, vol. 37, no. 2, pp. 177-218, 2007.

15 Switching boundary feedback stabilization for a star-shaped network of strings

Switching boundary feedback stabilization for a star-shaped network of strings

Martin Gugat
Friedrich-Alexander-Universität Erlangen-Nürnberg
Lehrstuhl 2 für Angewandte Mathematik
Martensstr. 3, 91058
Erlangen, Germany
Email: gugat@am.uni-erlangen.de

Mario Sigalotti
INRIA Nancy – Grand Est and
Institut Élie Cartan
UMR CNRS/INRIA/Nancy Université
BP 239, 54506
Vandœuvre-lès-Nancy, France
Email: mario.sigalotti@inria.fr

***Abstract*— We consider a star-shaped network with N strings that are coupled at one end and subject to a feedback control at the other end. Each feedback control is switched on and off by a time-dependent switching signal. We provide conditions on the switching laws that guarantee the exponential decay to zero of the system velocity. These sufficient conditions for stability ensure that at each moment in time only one wave arrives at the coupling note of the network or, alternatively, that at each moment in time only one wave is reflected at the controlled ends.**

I. Introduction

We consider a star-shaped network of N strings of finite length that are governed by the wave equation. At the boundary point zero the strings are coupled. At the other end of each string a feedback law is prescribed that requires the time derivative at this point to be proportional to the space derivative at this point. For a single string, this feedback law has been considered in [5], and it has been shown that the energy vanishes in finite time. In [7] it is shown that the result from [5] is stable in the sense that also with moving boundaries, the energy is driven to zero in finite time. In this paper we show that also on the network, the energy is driven to zero in finite time if the feedback control is active on all N boundary nodes for a sufficiently long time. Our particular interest in this paper is the question: What happens if at one of the nodes the feedback control becomes inactive? This need not be a fixed boundary node on the whole time interval but the inactivity may switch between different boundary nodes in time. The idea is that at each moment, it may happen that one of the N controlers is inactive, and we still want to have a stable system.

The boundary control of the wave equation has been studied by many authors (see e.g. [17], [16], [11], [12], [2], [20] and the references therein). A problem of optimal switching boundary control of a single string to rest in finite time has been considered in [8], where a single string with boundary control at both ends has been considered and, at each moment, at most one of the controls is allowed to be active. The corresponding problem for the heat equation has been analysed in [19] using an adapted adjoint calculus.

Networks of strings have been considered for example in [14], [15], [1], [13], [18] and an overview is given in [6]. In these works about networks, the nodes where a feedback control acts on the system are constant during the control process.

In contrast to this situation, in this paper we consider a system where these nodes may change as time proceeds. We are interested in the question: How many feedback controls must be switched on at each moment in time to achieve exponential decay? We show that $N-1$ controls are sufficient. It is essential that the choice of the inactive control need not be constant but can vary in a quite general way with time.

In a similar spirit, the notion on how often the control should be active in order to stabilize a string (with interior damping) was studied in [9]. Analogous questions have been addressed in [3], [4] for finite dimensional systems.

This paper has the following structure: First we define the problem of switching feedback boundary stabilization of a network of strings. Then we state our main results, which are two sufficient conditions for exponential decay of the derivatives in our system. First we state a backwards in time condition and then we state a forward condition.

For the proof of the backwards condition, we transform the initial boundary value problem in such a way that we can state it in terms of Riemann invariants. We show that if at each moment in time only one wave arrives at the coupling note of the network, the partial derivatives of the solution go to zero exponentially fast. This is the case if at each moment, at most one of the feedback nodes is switched off.

The proof of the forward condition is stated in the last part of the paper. It is based upon a Lyapunov function and thus uses a completely different method.

II. The System

Let $N \geq 3$ and consider N strings of length $L_i > 0$ ($i \in \{1, 2, \ldots, N\}$). Define $L = \max\{L_1, L_2, \ldots, L_N\}$. Let the corresponding wave speed $c > 0$ be given. For $i \in \{1, \ldots, N\}$ define the sets $\Omega_i = (0, \infty) \times (0,\ L_i)$. Define the set

$$B = \{(y_0^{(i)}, y_1^{(i)})_{i=1}^N :\ \partial_x y_0^{(i)} \in L^\infty(0, L_i),\ y_1 \in L^\infty(0, L_i),\ i \in \{1, ..., N\},\ y_0^{(i)}(0) = y_0^{(j)}(0),\ i, j \in i \in \{1, ..., N\}\}.$$

For $i \in \{1, \ldots, N\}$, let $\sigma_i : (0, \infty) \to \{0, 1\}$ be a measurable function. The equation $\sigma_i(t) = 0$ will indicate that

at time t the feedback at the end of string i is not active, whereas $\sigma_i(t) = 1$ means that the feedback is active.

For $(y_0^{(i)}, y_1^{(i)})_{i=1}^N \in B$ we consider the system (**S**) given by the equations

$$v^{(i)}(0,x) = y_0^{(i)}(x),\ x \in (0,L_i),\ i \in \{1,\dots,N\} \quad (1)$$

$$v_t^{(i)}(0,x) = y_1^{(i)}(x),\ x \in (0,L_i),\ i \in \{1,\dots,N\} \quad (2)$$

$$v_{tt}^{(i)}(t,x) = c^2\, v_{xx}^{(i)}(t,x),\ (t,x) \in \Omega_i,\ i \in \{1,\dots,N\} \quad (3)$$

$$v^{(i)}(t,0) = v^{(j)}(t,0),\ t \in (0,\infty), i,j \in \{1,\dots,N\} \quad (4)$$

$$v_x^{(1)}(t,0) + v_x^{(2)}(t,0) + \cdots + v_x^{(N)}(t,0) = 0,\ t \in (0,\infty) \quad (5)$$

$$c v_x^{(i)}(t,L_i) = -\sigma_i(t)\, v_t^{(i)}(t,L_i),\ t \in (0,\infty),\ i \in \{1,\dots,N\}. \quad (6)$$

III. Main Results

In this section we state the main results of this paper, which provide conditions on the switching functions σ_i that guarantee exponential decay of the derivatives in system (**S**).

Theorem 1: [Switching feedback stabilization of (**S**): Backward Condition] Consider system (**S**) defined in (1)–(6). Let $\lambda = L/c$.

If

$$\sum_{i=1}^N \sigma_i\left(t - \frac{L_i}{c}\right) \geq N-1 \text{ almost everywhere on } (0,\infty) \quad (7)$$

then the system state converges exponentially fast to a constant state, in the sense that for t almost everywhere on $(0,\infty)$ we have the inequality

$$\begin{aligned}&\operatorname{ess\,sup}\{|v_x^{(i)}(t,x)|,\ |v_t^{(i)}(t,x)| : x \in (0,L_i),\ i \in \{1,...,N\}\}\\ &\leq C \exp\left(\tfrac{\ln(f)}{2\lambda} t\right)\end{aligned} \quad (8)$$

where $f = \max\{\frac{2}{N},\ \frac{N-2}{N}\}$. The decay is uniform with respect to σ, that is the constant C in (8) is independent of the choice of σ verifying (7).

Proof. The proof is given in Section V.

Theorem 2: [Switching feedback stabilization of (**S**): Forward Condition] Consider system (**S**) defined in (1)–(6).

If

$$\sum_{i=1}^N \sigma_i\left(t + \frac{L_i}{c}\right) \geq N-1 \text{ almost everywhere on } (0,\infty) \quad (9)$$

then the energy of the state converges exponentially fast to zero, in the sense that

$$E(t) = \frac{1}{2}\sum_{i=1}^N \int_0^{L_i} \left(\frac{v_t^{(i)}(t,x)^2}{c^2} + v_x^{(i)}(t,x)^2\right) dx$$

satisfies

$$E(t) \leq C_1 \exp\left(-C_2 t\right) E(0), \quad (10)$$

for some $C_1, C_2 > 0$. The decay is uniform with respect to σ and the initial condition, that is, the constants C_1 and C_2 in (10) are independent of $(y_0^{(i)}, y_1^{(i)})_{i=1}^N$ and of the choice of σ verifying (9).

Proof. The proof is given in Section VI.

Remark. Conditions (7) and (9) in Theorems 1 and 2, respectively, cannot in general be relaxed by taking $N-2$ instead of $N-1$. Indeed, let us consider the case in which L_1 and L_2 are rationally dependent, that is, $L_1/L_2 \in \mathbf{Q}$. It is well known that there exist non-constant periodic solutions to the uncontrolled wave equation on the network (with Neumann boundary conditions) which are supported on the union of the first two strings. They are still solutions of (**S**), therefore, provided that $\sigma_1, \sigma_2 \equiv 0$. We proved the existence of solutions of (**S**) with $\sum_{i=1}^N \sigma_i(t \pm L_i/c) \equiv N-2$ that do not converge to any constant function.

IV. Transformation of the Problem

In this section we solve the initial boundary value problem (1)–(6) for given initial data. Let $(y_0^{(i)}, y_1^{(i)})_{i=1}^N \in B$ be given. We write the solution of the wave equation in the form of the d'Alembert solution

$$v^{(i)}(t,x) = \alpha_i(x+ct) + \beta_i(x-ct), \quad (11)$$

which means that we describe our solution in terms of the Riemann invariants or, in other words, as the sum of traveling waves. For an introduction to waves see [10]. For $i \in \{1,\dots,N\}$ the *initial conditions* (1), (2) yield the equations

$$y_0^{(i)}(x) = \alpha_i(x) + \beta_i(x),\ x \in (0,L_i) \quad (12)$$

$$y_1^{(i)}(x) = c\,[\alpha_i'(x) - \beta_i'(x)],\ x \in (0,L_i). \quad (13)$$

Hence we have

$$y_0^{(i)}(x) + (1/c)\int_0^x y_1^{(i)}(s)\,ds = 2\alpha_i(x) - k_1^{(i)}, \quad (14)$$

$$y_0^{(i)}(x) - (1/c)\int_0^x y_1^{(i)}(s)\,ds = 2\beta_i(x) + k_1^{(i)}, \quad (15)$$

for $x \in (0,L_i)$, where the real constants $k_1^{(i)}$ can be chosen as zero, which implies

$$\alpha_i(x) = \frac{1}{2}y_0^{(i)}(x) + \frac{1}{2c}\int_0^x y_1^{(i)}(s)\,ds, \quad (16)$$

$$\beta_i(x) = \frac{1}{2}y_0^{(i)}(x) - \frac{1}{2c}\int_0^x y_1^{(i)}(s)\,ds, \quad (17)$$

for $x \in (0,L_i)$. These representations imply that α_i', $\beta_i' \in L^\infty(0,L_i)$. We have shown that if (11) satisfies the initial conditions (1), (2) then with the normalization $k_1^{(i)} = 0$ (which is equivalent to $\alpha_i(0) = \beta_i(0)$) equations (16), (17) hold. The converse also holds: If α_i, β_i satisfy (16), (17), the initial conditions (1), (2) are valid for $v^{(i)}$ given by (11).

Writing the boundary condition

$$c v_x^{(i)}(t,L_i) = -\sigma_i(t)\, v_t^{(i)}(t,L_i)$$

in terms of α_i' and β_i' in the two cases $\sigma_i = 0$ and $\sigma_i = 1$, yields the equation

$$\alpha_i'(ct) = \left[\sigma_i\left(t - \frac{L_i}{c}\right) - 1\right]\beta_i'\left(2L_i - ct\right),\ t \in \left(\frac{L_i}{c}, \infty\right). \quad (18)$$

The node conditions (4), (5) imply that

$$\begin{pmatrix} \beta_1'(-ct) \\ \beta_2'(-ct) \\ \vdots \\ \beta_N'(-ct) \end{pmatrix} = A \begin{pmatrix} \alpha_1'(ct) \\ \alpha_2'(ct) \\ \vdots \\ \alpha_N'(ct) \end{pmatrix} \tag{19}$$

with the orthogonal symmetric reverberation matrix

$$A = \frac{N-2}{N} \begin{pmatrix} 1 & \frac{2}{2-N} & \frac{2}{2-N} & \cdots & \frac{2}{2-N} \\ \frac{2}{2-N} & 1 & \frac{2}{2-N} & \cdots & \frac{2}{2-N} \\ \vdots & & \ddots & & \vdots \\ \frac{2}{2-N} & \cdots & \frac{2}{2-N} & 1 & \frac{2}{2-N} \\ \frac{2}{2-N} & \frac{2}{2-N} & \cdots & \frac{2}{2-N} & 1 \end{pmatrix}. \tag{20}$$

Equations (18) and (19), together with the initial conditions (16) and (17), define α and β uniquely once the switching laws σ_i are given. Moreover, for what concerns their regularity, an induction argument shows that $\alpha_i' \in L^\infty(0,\infty)$ and $\beta_i' \in L^\infty(-\infty, L_i)$. Let us check that, for α and β constructed in this way, the functions $v^{(i)}$ obtained through (11) solve system (**S**). First we have

$$\begin{pmatrix} \frac{1}{c}v_t^{(1)}(t,0) \\ \frac{1}{c}v_t^{(2)}(t,0) \\ \vdots \\ \frac{1}{c}v_t^{(N)}(t,0) \end{pmatrix} = \begin{pmatrix} \alpha_1'(ct) - \beta_1'(-ct) \\ \alpha_2'(ct) - \beta_2'(-ct) \\ \vdots \\ \alpha_N'(ct) - \beta_N'(-ct) \end{pmatrix} =$$

$$= \frac{N-2}{N} \begin{pmatrix} \frac{2}{N-2} & \frac{2}{N-2} & \cdots & \frac{2}{N-2} \\ \vdots & \vdots & & \vdots \\ \frac{2}{N-2} & \frac{2}{N-2} & \cdots & \frac{2}{N-2} \end{pmatrix} \begin{pmatrix} \alpha_1'(ct) \\ \alpha_2'(ct) \\ \vdots \\ \alpha_N'(ct) \end{pmatrix},$$

which implies the equation $v_t^{(i)}(t,0) = v_t^{(j)}(t,0)$ for $t \in (0,\infty)$ and $i,j \in \{1,\dots,N\}$. Due to the definition of the set B, this implies that (4) is valid. Moreover, we have

$$\begin{pmatrix} v_x^{(1)}(t,0) \\ v_x^{(2)}(t,0) \\ \vdots \\ v_x^{(N)}(t,0) \end{pmatrix} = \begin{pmatrix} \alpha_1'(ct) + \beta_1'(-ct) \\ \alpha_2'(ct) + \beta_2'(-ct) \\ \vdots \\ \alpha_N'(ct) + \beta_N'(-ct) \end{pmatrix} =$$

$$= \begin{pmatrix} \frac{2N-2}{N} & -\frac{2}{N} & \cdots & -\frac{2}{N} \\ -\frac{2}{N} & \ddots & & \vdots \\ \vdots & & \ddots & -\frac{2}{N} \\ -\frac{2}{N} & \cdots & -\frac{2}{N} & \frac{2N-2}{N} \end{pmatrix} \begin{pmatrix} \alpha_1'(ct) \\ \alpha_2'(ct) \\ \vdots \\ \alpha_N'(ct) \end{pmatrix},$$

which implies the equation $v_x^{(1)}(t,0) + v_x^{(2)}(t,0) + \cdots + v_x^{(N)}(t,0) = 0$ for $t \in (0,\infty)$, hence (5) holds.

V. Proof of Theorem 1

We start the proof of Theorem 1 by the following auxiliary result.

Lemma 3: Assume that condition (7) is satisfied. Then the following inequality holds for all natural numbers k:

$$\begin{aligned} &\operatorname{ess\,sup}_{s>2kL} \max_{i\in\{1,\dots,N\}} \{|\alpha_i'(s)|,\ |\beta_i'(-s)|\} \\ &\quad\le\ \sqrt{N}\, f^k \operatorname{ess\,sup}_{s\in(0,2L)} \max_{i\in\{1,\dots,N\}} |\alpha_i'(s)| \end{aligned} \tag{21}$$

where $f = \max\{\frac{2}{N},\ \frac{N-2}{N}\}$.

Moreover, for all $t \ge (2k+1)\lambda$ the following inequality holds:

$$\begin{aligned} &\operatorname{ess\,sup}\{|v_x^{(i)}(t,x)|, |v_t^{(i)}(t,x)| : x \in (0,L_i),\ i \in \{1,...,N\}\} \\ &\le \max\{1,\ c\}\Big(1+\sqrt{N}\Big) f^k \max_{j\in\{1,\dots,N\}} \|\alpha_j'(s)\|_{L^\infty(0,2L)}. \end{aligned} \tag{22}$$

Proof. The idea of the proof is that for all $s > 2kL$ we can go backwards in $(0,s)$ until a point in the interval $(0,2L)$ is reached in at least k steps that are less than or equal to $2L$. In each of these steps, the essential supremum is reduced at least by a factor f.

This can be seen in the following way. Condition (7) implies that there exists at most one number $k \in \{1,\dots,N\}$ with $\alpha_k'(ct) \neq 0$. For the other $N-1$ dervivatives we have $\alpha_j'(ct) = 0$, $j \neq k$. Due to (18), we have the inequality

$$|\alpha_k'(ct)| \le |\beta_k'(-c(t-2L_k/c))|.$$

Due to (19) we have

$$\begin{pmatrix} \beta_1'(-ct+2L_k) \\ \beta_2'(-ct+2L_k) \\ \vdots \\ \beta_N'(-ct+2L_k) \end{pmatrix} = Aw.$$

If we apply (18) for the time $\tilde{t} = t - 2L_k/c$, we see that also the vector

$$w = \begin{pmatrix} \alpha_1'(ct-2L_k) \\ \alpha_2'(ct-2L_k) \\ \vdots \\ \alpha_N'(ct-2L_k) \end{pmatrix}$$

can have at most one nonzero component.

On account of the definition (20) of the matrix A this yields the inequality

$$|\beta_k'(-ct+2L_k)| \le f \max_{l\in\{1,\dots,N\}} |\alpha_l'(ct-2L_k)|.$$

Hence we have the inequality

$$\begin{aligned} \max_{l\in\{1,\dots,N\}} |\alpha_l'(ct)| &= |\alpha_k'(ct)| \le |\beta_k'(-c(t-2L_k/c))| \\ &\le f \max_{l\in\{1,\dots,N\}} |\alpha_l'(ct-2L_k)|. \end{aligned}$$

For s_0 almost everywhere in $(2L,\infty)$ this yields the inequality

$$\max_{i\in\{1,\dots,N\}} \{|\alpha_i'(s_0)|\} \le f \max_{i\in\{1,\dots,N\}} \|\alpha_i'(s)\|_{L^\infty(0,s_0-\hat{L})},$$

where

$$\hat{L} = \min\{L_1,\dots,L_N\}.$$

By induction for $|\alpha_i'(s)|$ we obtain the inequality

$$\operatorname*{ess\,sup}_{s>2kL} \max_{i\in\{1,\dots,N\}} \{|\alpha_i'(s)|\} \le f^k \max_{i\in\{1,\dots,N\}} \|\alpha_i'(s)\|_{L^\infty(0,2L)}. \tag{23}$$

Due to (19) and the fact that the matrix A is orthogonal we have

$$\begin{aligned}\max_{l\in\{1,\dots,N\}} |\beta_l'(-s)| &\le \sqrt{\sum_{l=1}^N |\beta_l'(-s)|^2} = \sqrt{\sum_{l=1}^N |\alpha_l'(s)|^2} \\ &\le \sqrt{N} \max_{l\in\{1,\dots,N\}} |\alpha_l'(s)|\end{aligned}$$

hence also the inequality (21) including the factor $\sqrt{N}$ for $|\beta_i'(s)|$ follows.

If $x+ct \ge 2kL$ and $x-ct \le -2kL$ for all $i \in \{1,2,..,N\}$ we have

$$\begin{aligned}&\max\{|v_x^{(i)}(t,x)|,\ |v_t^{(i)}(t,x)|\} \\ &\le \max\{1,\ c\}\,[|\alpha_i'(x+ct)| + |\beta_i'(x-ct)|] \\ &\le \max\{1,\ c\}\left(1+\sqrt{N}\right) f^k \operatorname*{ess\,sup}_{s\in(0,2L)} \max_{j\in\{1,\dots,N\}} |\alpha_j'(s)|.\end{aligned}$$

This implies inequality (22). □

In order to complete the proof of Theorem 1 let us define $\rho(t)$ as

$$\operatorname{ess\,sup}\{|v_x^{(i)}(t,x)|,\ |v_t^{(i)}(t,x)| : x \in (0,L_i),\ i \in \{1,\dots,N\}\}.$$

Due to (**??**) we have

$$\begin{aligned}\rho(t) \le\ & \max\{1,\ c\} \operatorname{ess\,sup}\{|\alpha_i'(x+ct)| + |\beta_i'(x-ct)| : \\ & x \in (0,L_i),\ i \in \{1,\dots,N\}\}.\end{aligned}$$

We claim that ρ satisfies

$$\rho(t) \le C_0, \qquad t \in (0,3\lambda), \tag{24}$$

$$\rho(t) \le C_1 f^k, \quad t > (2k+1)\lambda, \quad k \in \mathbf{N} \tag{25}$$

with C_0 and C_1 only depending on the initial condition $(y_0^{(i)}, y_1^{(i)})_{i=1}^N$ and not on the choice of σ verifying (7).

For $x \in (0,L_i)$, the values of $\alpha_i(x)$ are given by (16) and the values of $\beta_i(x)$ are given by (17). Hence the constant

$$\begin{aligned}C_* = \ & 2\max\{1,\ c\} \operatorname{ess\,sup}\{|\alpha_i'(x)|,\ |\beta_i'(x)| : \\ & x \in (0,L_i),\ i \in \{1,\dots,N\}\}\end{aligned}$$

only depends on the initial condition and satisfies $\rho(0) \le C_*$. For $x \in (L_i, 2L_i)$, equation (18) yields the inequality

$$|\alpha_i'(x)| \le |\beta_i'(2L_i - x)| \le C_*/2, \quad x \in (L_i, 2L_i).$$

Hence for all $x \in (0,2L_i)$ we have $|\alpha_i'(x)| \le C_*/2$. Since A has the matrix norm $\|A\|_\infty = 3 - \frac{4}{N}$, the node condition (19) implies that

$$|\beta_i'(-ct)| \le \left(3 - \frac{4}{N}\right) \max\{|\alpha_j'(ct)| : j = 1,\dots,N\} \tag{26}$$

for $t \in (0,\infty)$.

Moreover, due to (18) we have

$$|\alpha_i'(x)| \le \left(3 - \frac{4}{N}\right) \max\{|\alpha_j'(x-2L_i)| : j = 1,\dots,N\}$$

for $x > 2L_i$. Hence, by recurrence, if $x < 2k\hat{L}$ for some integer k, then

$$|\alpha_i'(x)| \le \frac{C_*}{2}\left(3 - \frac{4}{N}\right)^{k-1}.$$

As a consequence, if $t < 2k\hat{L}/c$ and $x \in (0, L_i)$, then

$$\begin{aligned}|\alpha_i'(ct+x)| &\le \frac{C_*}{2}\left(3 - \frac{4}{N}\right)^k, \\ |\beta_i'(x-ct)| &\le \frac{C_*}{2}\left(3 - \frac{4}{N}\right)^k,\end{aligned}$$

where the second inequality uses (26). It follows that

$$\rho(t) \le \max\{1,c\} C_* \left(3 - \frac{4}{N}\right)^k, \quad \text{if } t < 2k\hat{L}/c.$$

Now we choose $\bar{k}$ such that $3L < 2\bar{k}\hat{L}$ and set

$$C_0 = \max\{1,c\} C_* \left(3 - \frac{4}{N}\right)^{\bar{k}}.$$

If $t > (2k+1)\lambda$ for some integer k inequality (22) implies

$$\begin{aligned}\rho(t) &\le \max\{1,c\}\left(1+\sqrt{N}\right) \max_{j\in\{1,\dots,N\}} \|\alpha_j'(s)\|_{L^\infty(0,2L)}\ f^k \\ &\le \max\{1,\ c\}\left(1+\sqrt{N}\right)\ C_0\, f^k.\end{aligned}$$

Hence we choose

$$C_1 = \max\{1,\ c\}\left(1+\sqrt{N}\right)\ C_0.$$

Since $\rho(t)$ satisfies (24) and (25), it can be bounded from above by $f^{-2}\max\{C_0, C_1\} \exp\left(\frac{\ln(f)}{2\lambda} t\right)$. □

VI. Proof of Theorem 2

Let $\varepsilon > 0$ and define, for $i = 1,\dots,N$,

$$\Xi_i = \left\{ x \in \mathbf{R}^N :\ \sum_{j\ne i} x_j^2 < \varepsilon x_i^2 \right\}.$$

Hence, Ξ_i is a cone with axial symmetry with respect to the axis spanned by the i-th vector of the canonical basis of $\mathbf{R}^N$. Let $\Xi = \cup_{i=1}^N \Xi_i$.

Lemma 4: There exists $k_1 > 0$ depending only on ε such that, if x belongs to $\mathbf{R}^N \setminus \Xi$, then

$$\min_{i=1,\dots,N} \sum_{j\ne i} x_j^2 \ge k_1 \|x\|^2. \tag{27}$$

Proof. Let i be the index achieving the minimization in (27). Since $x \notin \Xi_i$, then

$$\sum_{j\ne i} x_j^2 \ge \frac{1}{2}\sum_{j\ne i} x_j^2 + \frac{\varepsilon x_i^2}{2} \ge \min\left(\frac{1}{2}, \frac{\varepsilon}{2}\right) \|x\|^2$$

and the lemma is proved. □

Define Θ as the set of non-negative times t such that $\beta'(-ct) \in \Xi$.

Lemma 5: Let $k_2 \in (0, \min\{2, N-2\}/N)$. Then, for every $\varepsilon > 0$ small enough, almost every $t \in \Theta$ and every $i = 1, \dots, N$,

$$|\alpha_i'(ct)| \geq k_2 \|\alpha'(-ct)\|. \tag{28}$$

Moreover, for almost every $t \in \{\tau \mid \tau \geq 4\lambda,\ \tau \in \Theta\}$, either $\alpha'(ct) = 0$ or $t - 2L_i \in \Theta$ for at most one $i \in \{1, \dots, N\}$ and $\sigma_l(t - L_l/c) = 0$ for every $l \in \{1, \dots, N\}$.

Proof. Let

$$\Upsilon = \{x \in \mathbf{R}^N : |x_i| \geq k_2 \|x\| \text{ for every } i = 1, \dots, N\}.$$

In order to prove the first part of the statement, we have to show that for every $\varepsilon > 0$ small enough and almost every $t \in \Theta$, $\alpha'(ct) \in \Upsilon$.

Since A is idempotent and because of (19), $\alpha'(ct) \in A\Xi$ for almost every $t \in \Theta$. Notice that $A\Xi$ is the union of the N cones with axial symmetry with respect to the columns of A and with the same aperture as the Ξ_i's.

We have to show that for every $\varepsilon > 0$ small enough $A\Xi$ is contained in Υ. It suffices to notice that the boundary of Υ is invariant by multiplication by a scalar and that the each vector corresponding to a column of A is in the interior of Υ. (Indeed, if x is a column of A, then $\|x\| = 1$ and $|x_i| = (N-2)/N$ or $|x_i| = 2/N$.) Then for ε small enough every vector of $A\Xi \backslash \{0\}$ belongs to the interior of Υ.

As for the second part of the statement, take $t \in \Theta$ such that $t \geq 4\lambda$ and $\alpha'(ct) \neq 0$. Because of the first part of the lemma, $\alpha_l'(ct) \neq 0$ for every $l \in \{1, \dots, N\}$. It follows from (18) that $\beta_l'(2L_l - ct) \neq 0$ and $\sigma_l'(2L_l - ct) = 0$ for every $l \in \{1, \dots, N\}$.

Assume now by contradiction that $t - 2L_i/c, t - 2L_j/c \in \Theta$ with $i \neq j$. In particular, both $\beta_i'(2L_i - ct)$ and $\beta_j'(2L_j - ct)$ are nonzero. Hence, $\alpha'(ct - 2L_i)$ and $\alpha'(ct - 2L_j)$ are nonzero. Since $t - 2L_i/c, t - 2L_j/c \in \Theta$, we can apply again the first part of the lemma and (18), deducing that $\sigma_i(t - (2L_j + L_i)/c) = \sigma_j(t - (2L_i + L_j)/c) = 0$.

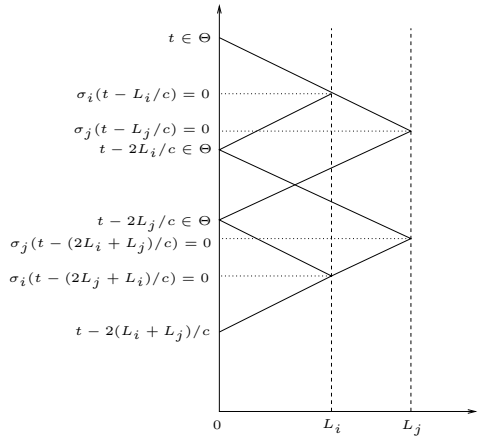


Fig. 1. Contradiction argument for $t - 2L_i/c, t - 2L_j/c \in \Theta$.

This contradicts condition (9) when we take as t the time $t - 2(L_i + L_j)/c$ and concludes the proof of the lemma. □

Let us complete the proof of Theorem 2.

The time-derivative of the energy $E(t)$ can be computed using equations (3)–(6) and is given by

$$\dot{E}(t) = -\frac{1}{c} \sum_{i=1}^{N} \sigma_i(t) \left(v_t^{(i)}(t, L_i)\right)^2$$

for almost every t. In particular, E is non-increasing. Notice that $v_t^{(i)}(L_i) = c(\alpha_i'(L_i + ct) - \beta_i'(L_i - ct))$ (see (11)) and that $\alpha_i'(L_i + ct) = 0$ if $\sigma_i(t) = 1$ (see (18)). Therefore,

$$\dot{E}(t) = -c \sum_{i=1}^{N} \sigma_i(t) \beta_i(L_i - ct)^2.$$

Let

$$F(t) = \sum_{i=1}^{N} E\left(t + \frac{L_i}{c}\right).$$

Then

$$\dot{F}(t) \leq -c \sum_{i=1}^{N} \sigma_i \left(t + \frac{L_i}{c}\right) \beta_i(-ct)^2.$$

Lemma 4 and condition (9) guarantee that if $t \notin \Theta$, then

$$\dot{F}(t) \leq -ck_1 \|\alpha'(ct)\|^2.$$

On the other hand, according to Lemma 5, for almost every $t \in \Theta$ such that $t \geq 4\lambda$, either $\alpha'(ct) = 0$ or for all but possibly one $i \in \{1, \dots, N\}$ we have

$$\begin{aligned} \dot{F}(t - 2L_i/c) &\leq -ck_1 \|\alpha'(ct - 2L_i)\|^2 \\ &= -ck_1 \|\beta'(2L_i - ct)\|^2 \\ &\leq -ck_1 |\beta_i'(2L_i - ct)|^2 \\ &= -ck_1 |\alpha_i'(ct)|^2 \ \leq\ -ck_1 k_2^2 \|\alpha'(ct)\|^2. \end{aligned}$$

Let $G(t) = F(t) + \sum_{i=1}^{N} F(t - 2L_i/c)$. Then, for almost every $t \geq 4\lambda$,

$$\dot{G}(t) \leq -(N-1)ck_1 k_2^2 \|\alpha'(ct)\|^2, \tag{29}$$

where we used the inequality $(N-1)k_2^2 \leq 1$.

Notice that

$$\begin{aligned} &2(\alpha_i'(x + ct)^2 + \beta_i'(x - ct)^2) = \\ &= (\alpha_i'(x + ct) + \beta_i'(x - ct))^2 + (\alpha_i'(x + ct) - \beta_i'(x - ct))^2 \\ &= \tfrac{v_t^{(i)}(t,x)^2}{c^2} + v_x^{(i)}(t, x)^2. \end{aligned}$$

Thus we have

$$\begin{aligned} E(t) &\leq \int_0^L \left(\|\alpha'(x + ct)\|^2 + \|\beta'(x - ct)\|^2\right) dx \\ &= \int_0^L \left(\|\alpha'(x + ct)\|^2 + \|\alpha'(ct - x)\|^2\right) dx \\ &= c \int_{t-\lambda}^{t+\lambda} \|\alpha'(cs)\|^2 ds. \end{aligned}$$

If $t \geq 5\lambda$ this and (29) imply the inequality

$$G(t + \lambda) - G(t - \lambda) \leq -(N-1)k_1 k_2^2 E(t).$$

By monotonicity of E and definition of G,

$$G(t) \leq (N+1)F(t - 2\lambda) \leq N(N+1)E(t - 2\lambda)$$

so that

$$G(t+\lambda) - G(t-\lambda) \leq -\frac{N-1}{N(N+1)} k_1 k_2^2 G(t+2\lambda).$$

Hence $G(t)$ decays exponentially to zero as t goes to infinity. Moreover, $G(5\lambda) \leq N(N+1)E(0)$. Since k_1, k_2, N and λ do not depend on the initial conditions nor on σ, we have that $G(t) \leq C_1 \exp(-C_2 t)\, E(0)$ with C_1 and C_2 independent of $(y_0^{(i)}, y_1^{(i)})_{i=1}^N$ and of the choice of σ verifying (9).

Inequality (10) follows and this concludes the proof of Theorem 2. □

VII. Conclusion

For a single string it is well known that a velocity feedback at one end with a special feedback parameter steers the solution to a constant state in finite time; the semigroup describing the corresponding solution is nilpotent. In this paper we prove that a similar situation occurs for star-shaped networks with boundary feedback at all boundary nodes: The partial derivatives of the solution vanish after finite time and the system state becomes constant.

If the feedback is switched off at one of the boundary nodes in such a way that at each moment of the delayed time axes shifted by the corresponding travel time of the signals to the coupling node at least $N-1$ of the feedback controls are active, then the partial derivatives of the solution decay exponentially fast. Note that the node where the control is switched off need not be constant but can vary with time.

This result may be interpreted in the following way: In the boundary feedback stabilization of a star-shaped network the exponential decay property is not destroyed if at each moment one of the feedback controllers does not work.

It is an open question whether the corresponding result holds for a tree-shaped network of strings with control at all boundary nodes. For the analysis of this problem, the most promising approach appears to be the consideration of the dynamics of the energy on the whole tree.

Acknowledgment

This paper was supported by the PROCOPE program of DAAD, D/0811409 and Égide, 20161VM. It took benefit from discussions during the meeting *Partial differential equations, optimal design and numerics 2009* at the Benasque Center for Science Pedro Pascual.

References

[1] K. Ammari, M. Jellouli, and M. Khenissi. Stabilization of generic trees of strings. *J. Dyn. Control Syst.*, 11(2):177–193, 2005.

[2] S. A. Avdonin and S. A. Ivanov. *Families of exponentials*. Cambridge University Press, Cambridge, 1995.

[3] A. Chaillet, Y. Chitour, A. Loría, and M. Sigalotti. Uniform stabilization for linear systems with persistency of excitation: the neutrally stable and the double integrator cases. *Math. Control Signals Systems*, 20(2):135–156, 2008.

[4] Y. Chitour and M. Sigalotti. On the stabilization of persistently excited linear systems. *Preprint*, 2009.

[5] S. Cox and E. Zuazua. The rate at which energy decays in a string damped at one end. *Indiana Univ. Math. J.*, 44(2):545–573, 1995.

[6] R. Dáger and E. Zuazua. *Wave propagation, observation and control in 1-d flexible multi-structures*, volume 50 of *Mathématiques & Applications (Berlin) [Mathematics & Applications]*. Springer-Verlag, Berlin, 2006.

[7] M. Gugat. Optimal boundary feedback stabilization of a string with moving boundary. *IMA J. Math. Control Inform.*, 25(1):111–121, 2008.

[8] M. Gugat. Optimal switching boundary control of a string to rest in finite time. *ZAMM Z. Angew. Math. Mech.*, 88(4):283–305, 2008.

[9] A. Haraux, P. Martinez, and J. Vancostenoble. Asymptotic stability for intermittently controlled second-order evolution equations. *SIAM J. Control Optim.*, 43(6):2089–2108, 2005.

[10] R. Knobel. *An introduction to the mathematical theory of waves*, volume 3 of *Student Mathematical Library*. American Mathematical Society, Providence, RI, 2000. IAS/Park City Mathematical Subseries.

[11] W. Krabs. Optimal control of processes governed by partial differential equations. II. Vibrations. *Z. Oper. Res. Ser. A-B*, 26(3):63–86, 1982.

[12] W. Krabs. *On moment theory and controllability of one-dimensional vibrating systems and heating processes*, volume 173 of *Lecture Notes in Control and Information Sciences*. Springer-Verlag, Berlin, 1992.

[13] J. E. Lagnese, G. Leugering, and E. J. P. G. Schmidt. *Modeling, analysis and control of dynamic elastic multi-link structures*. Systems & Control: Foundations & Applications. Birkhäuser Boston Inc., Boston, MA, 1994.

[14] G. Leugering. On feedback controls for dynamic networks of strings and beams and their numerical simulation. In *Control and estimation of distributed parameter systems: nonlinear phenomena (Vorau, 1993)*, volume 118 of *Internat. Ser. Numer. Math.*, pages 301–326. Birkhäuser, Basel, 1994.

[15] G. Leugering and E. Zuazua. On exact controllability of generic trees. In *Contrôle des systèmes gouvernés par des équations aux dérivées partielles (Nancy, 1999)*, volume 8 of *ESAIM Proc.*, pages 95–105. Soc. Math. Appl. Indust., Paris, 2000.

[16] J.-L. Lions. Exact controllability, stabilization and perturbations for distributed systems. *SIAM Rev.*, 30(1):1–68, 1988.

[17] D. L. Russell. Nonharmonic Fourier series in the control theory of distributed parameter systems. *J. Math. Anal. Appl.*, 18:542–560, 1967.

[18] J. Valein and E. Zuazua. Stabilization of the wave equation on 1-d networks. *SIAM J. Contr. Optim*, 48(4):2771–2797, 2009.

[19] E. Zuazua. Switching controls. *Journal of the European Mathematical Society,* to appear.

[20] E. Zuazua. Optimal and approximate control of finite-difference approximation schemes for the 1D wave equation. *Rend. Mat. Appl. (7)*, 24(2):201–237, 2004.

16 Singular Perturbation Approach to Pulse-Width Modulated Control in Nonlinear Dynamical Systems

Singular Perturbation Approach to Pulse-Width Modulated Control in Nonlinear Dynamical Systems

Valery D. Yurkevich
Automation Department
Novosibirsk State Technical University
Novosibirsk, Russia, 630092
Email: yurkev@ac.cs.nstu.ru

***Abstract*—The problem of robust controller design under PWM feedback is discussed in terms of Filippov's average model where control variable is a duty ratio function. The proposed controller is an extension of PI/PID control scheme under PWM feedback. The presented design methodology guarantees desired output transient performance indices by inducing of two-time-scale motions in the closed-loop system. Stability conditions imposed on the fast and slow modes and sufficiently large mode separation rate between fast and slow modes can ensure that the full-order closed-loop nonlinear system achieves the desired properties in such a way that the output transient performances are desired and insensitive to external disturbances and plant's parameter variations. The method of singular perturbations is used throughout the paper in order to get explicit expressions for evaluation of the controller parameters. Simulation results of tracking control for magnetic levitation system are presented as an example of the application for the discussed PWM control design methodology.**

I. Introduction

There are many control systems with pulse-width modulation (PWM) in feedback loop that are widely used in applications where the most important ones are power converters and motor control systems [1]-[4]. The great improvement of power electronic switching device characteristics and the drastic decreasing of switching device production cost give a possibility to provide a high sampling frequency in PWM control systems. Therefore, among various PWM techniques the principle of equivalent areas [5] is efficiently used where the continuous-time or discrete-time control algorithms are designed at the beginning, and then ones should be re-designed in order to be implemented in a PWM feedback loop [6].

Theoretical problems of stability and oscillations for nonlinear pulse-modulated systems are investigated based on averaging method and Lyapunov function method with help of V.A.Yakubovich's frequency theorem in [7], [8].

The existence of an equivalence between sliding modes of variable structure control and PWM control responses under the high frequency sampling gives other possibilities for PWM controller design [9]. It was shown in [10], if PWM controller is not saturated and the sampling frequency tends to infinity, then the response of discontinuously controlled system coincides with Filippov's average model [11] where control variable is represented by duty ratio function. Hence, various type of continuous-time or discrete-time controllers can be designed based on the Filippov's average model, for instance, proportional-integral (PI) or proportional-integral-derivative (PID) controllers.

Problems of PI (PID) control system analysis and design are treated in a huge set of publications, for instance, [12], [13], that are only few ones. For example, the well known Ziegel-Nichols tuning rules [14] or its various modifications are widely used for selection of controller parameters. In the presence of plant uncertainty, in order to fetch out the best PI and PID controllers in accordance with the assigned design objectives, a set of tuning rules, identification and adaptation schemes has been developed [15], [16]. The main disadvantage for the most part of the existing procedures on PI or PID controller design is that the desired transient performances can not be guaranteed in the presence of nonlinear plant parameter variations and unknown external disturbances.

The objectives of this paper are the analysis and design of control systems with PWM feedback loop for nonlinear plant model in the presence of plant's parameter variations, and unknown external disturbances. The discussed approach to controller design is based on the design methodology presented in [17], that guarantees desired output transients by inducing of two-time-scale motions in the closed-loop system. Stability conditions imposed on the fast and slow modes and sufficiently large mode separation rate between fast and slow modes can ensure that the full-order closed-loop nonlinear system achieves the desired properties in such a way that the output transient performances are desired and insensitive to external disturbances and plant's parameter variations. The stability of fast-motion transients in the closed-loop system is provided by proper selection of controller parameters, as well as slow-motion transients correspond to the stable reference model of desired mapping from reference input into controlled output. The method of singular perturbations [18]-[23] is used throughout the paper.

The paper is a continuation of [24] and one is organized as follows. First, the control design objectives are stated. Second, the Filippov's average model for nonlinear nonaffine-in-control system under high-frequency sampling is introduced. Third, the generalized structure of feedback controller, which is called as a universal controller, for nonlinear systems of an arbitrary degree is presented where the proposed controller is an extension of PI/PID control scheme. Fourth, the two-time-scale motions analysis of the closed-loop system properties is discussed. Such

questions as conditions for not saturated behavior of pulse-width modulator and attenuation of high-frequency chattering caused by the switching mode of the pulse-width modulator are addressed as well. Finally, simulation results of tracking control for magnetic levitation system are presented as an example of the application for the proposed PWM control design methodology.

II. Control Problem Statement

Consider the SISO nonlinear system of the following form:

$$x^{(n)} = f(X, w, u) \tag{1}$$

where $X = [x, x^{(1)}, \ldots, x^{(n-1)}]^T$ is the state vector and x is the measurable output of the system (1).

It is assumed that $f(X, w, u)$ is an unknown scalar continuous bounded function of X, w, and u on a bounded set $\Omega_{X,w,u} := \Omega_X \times \Omega_w \times \Omega_u$.

Remark 1: The discussed nonlinear system given by (1), in general, can be nonaffine-in-control system, that is an explicit inversion of the function $z = f(X, w, u)$ with respect to control variable u is impossible for given z, x, and w. For instance, the system given by $\dot{x} = x^3 + u(1 - u^2)$ is the non-affine in control one.

Let $u(t)$ is the scalar control variable, where u takes one of two possible values in $\Omega_u := \{u^-, u^+\}$.

Assume that the pulse-width modulated control for the system (1) is defined as the switching function $u(t)$ given by

$$u = \begin{cases} u^+ & \text{for } \ t_\kappa < t \le t_\kappa + \chi(t_\kappa)T_s \\ u^- & \text{for } \ t_\kappa + \chi(t_\kappa)T_s < t \le t_\kappa + T_s \end{cases} \tag{2}$$

where T_s is the sampling period of the pulse-width modulation, χ is the duty ratio function which takes values in the interval $[0, 1]$, $\chi(t_\kappa)$ is the duty ratio function at the time instance t_κ, $t_\kappa = \kappa T_s$, and $\kappa = 0, 1, 2, \ldots$.

A control system is being designed so that the condition

$$\lim_{t \to \infty} x(t) = r \tag{3}$$

holds, where $r = const$. Moreover, the output transients of $x(t)$ should have the desired performance indices. These transients should not depend on nonlinearity of the system (1), the external disturbance or varying parameter represented by $w(t)$.

The block diagram of the discussed control system is shown in Fig. 1, where the plant (P) is represented by (1), the pulse-width modulator (PWM) is given by (2), and the controller (C) should by designed later on.

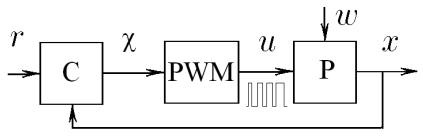


Fig. 1. Block diagram of the control system.

III. Control Via Time-Scale Separation

A. Filippov's Average Model

Denote

$$f^+(X, w) := f(X, w, u^+), \quad f^-(X, w) := f(X, w, u^-).$$

Then the system (1)-(2) can be rewritten as

$$x^{(n)} = \nu f^+(X, w) + (1 - \nu) f^-(X, w) \tag{4}$$

where ν is defined as the ideal switching function

$$\nu = \begin{cases} 1 & \text{for } \ t_\kappa < t \le t_\kappa + \chi(t_\kappa)T_s \\ 0 & \text{for } \ t_\kappa + \chi(t_\kappa)T_s < t \le t_\kappa + T_s \end{cases} \tag{5}$$

Assumption 1: The pulse-width modulator given by (2) is not saturated, that is the following condition $0 < \chi < 1$ holds.

Assumption 2: The sampling period T_s is assumed to be sufficiently small in comparison with time constants associated with the dynamics of the system (1).

In accordance with Assumptions 1 and 2, by following to the Filippov's approach [11], the geometric approach to PWM control [9], and Theorem A.1 in the paper by [10], the response of discontinuously controlled system given by (1) and (2) coincides with Filippov's average model

$$x^{(n)} = \chi f^+(X, w) + (1 - \chi) f^-(X, w) \tag{6}$$

where χ is the duty ratio and $\chi \in (0, 1)$ by Assumption 1.

The closed-loop system properties will be treated below based on the average model (6), that is rewritten as

$$x^{(n)} = f^-(X, w) + [f^+(X, w) - f^-(X, w)]\chi. \tag{7}$$

Remark 2: The nonlinear non-affine in control system (1) with the high-frequency pulse-width modulator (2) yields the average model (7) which is affine in control system.

B. Control Law

By following to the design methodology [17], [24], consider the universal controller given by

$$\begin{aligned} \mu^q \chi^{(q)} + d_{q-1}\mu^{q-1}\chi^{(q-1)} + \cdots + d_1 \mu \chi^{(1)} \\ = k[F(X, r) - x^{(n)}] \end{aligned} \tag{8}$$

where μ is the small positive parameter, $q \ge n$, and

$$F(X, r) = -\frac{a^0_{n-1}}{T} x^{(n-1)} - \cdots - \frac{a^0_1}{T^{n-1}} x^{(1)} + \frac{1}{T^n}[r - x].$$

The controller parameters $a^0_{n-1}, \ldots, a^0_1, T$ are selected such that the polynomial

$$T^n s^n + a^0_{n-1} T^{n-1} s^{n-1} + \cdots + a^0_2 T^2 s^2 + a^0_1 T s + 1 \tag{9}$$

has the desired root distribution inside the left part of the s-plane, where roots of the polynomial (9) are defined by the requirements imposed on the desired output transient performance indices of $x(t)$ in the system (1).

Remark 3: The control law (8) can be expressed in terms of transfer functions, that is

$$\chi(s) = \frac{k}{\mu(\mu^{q-1}s^{q-1} + d_{q-1}\mu^{q-2}s^{q-2} + \cdots + d_2\mu s + d_1)} \times\left\{\frac{[r(s)-x(s)]}{sT^n} - (s^{n-1}+\cdots+\frac{a_2^0 s}{T^{n-2}}+\frac{a_1^0}{T^{n-1}})x(s)\right\}$$

where this controller is proper and one is implemented without an ideal differentiation of $x(t)$ or $r(t)$ due to $q \geq n$.

Remark 4: The conventional PI controller results from (8) when $q = n = 1$. The PI controller with an additional lowpass filtering results from (8) when $q > n = 1$. The proper PID controller with an additional lowpass filtering results from (8) when $q > n = 2$.

Remark 5: If $q \geq n$, then the controller (8) can be rewritten as the system of state space differential equations given by

$$\begin{aligned} \dot{U}_\chi &= A_c U_\chi + B_c x + E_c r \\ \chi &= C_c U_\chi + D_c x \end{aligned} \tag{10}$$

where $U_\chi \in \mathbb{R}^q$. Note, the relation between the duty ratio function $\chi(t)$ and the control variable $u(t)$ is defined by the switching function (2).

C. Time-Scale Separation

Consider the closed-loop system equations given by the average model (7) and controller (8), that are

$$\begin{aligned} x^{(n)} &= f^-(X,w) + [f^+(X,w) - f^-(X,w)]\chi \\ \mu^q \chi^{(q)} &+ d_{q-1}\mu^{q-1}\chi^{(q-1)} + \cdots + d_1\mu\chi^{(1)} \\ &= k[F(X,r) - x^{(n)}]. \end{aligned}$$

The replacement of $x^{(n)}$ in (8) by the right member of (7) yields the closed-loop system equations in the form

$$\begin{aligned} x^{(n)} &= f^-(X,w) + [f^+(X,w) - f^-(X,w)]\chi \\ \mu^q \chi^{(q)} &+ d_{q-1}\mu^{q-1}\chi^{(q-1)} + \cdots + d_1\mu\chi^{(1)} \\ +k[f^+(X,w) &- f^-(X,w)]\chi = k[F(X,r) - f^-(X,w)]. \end{aligned} \tag{11}$$

Denote $x_1 = x$, $x_2 = x^{(1)}$, ..., $x_n = x^{(n-1)}$, $\chi_1 = \chi$, $\chi_2 = \mu\chi^{(1)}$, ..., $\chi_q = \mu^{q-1}\chi^{(q-1)}$. Then the closed-loop system (11) may be rewritten as the following system of singularly perturbed differential equations:

$$\begin{aligned} \dot{x}_i &= x_{i+1},\ i = 1,2,\ldots,n-1 \\ \dot{x}_n &= f^-(X,w) + [f^+(X,w) - f^-(X,w)]\chi_1 \\ \mu\dot{\chi}_j &= \chi_{j+1},\ j = 1,2,\ldots,q-1 \\ \mu\dot{\chi}_q &= -k[f^+(X,w) - f^-(X,w)]\chi_1 - d_1\chi_2 \\ &\quad -\cdots - d_{q-1}\chi_q + k[F(X,r) - f^-(X,w)] \end{aligned} \tag{12}$$

where two-time-scale motions are forced as $\mu \to 0$. Hence, fast and slow modes are artificially forced in the closed-loop system and the time-scale separation between these modes depends on the parameter μ.

In order to enable usage of the standard technique for two-time-scale motions analysis, take $t = \mu t_0$. Hence, from (12), the system

$$\begin{aligned} \frac{dx_i}{dt_0} &= \mu x_{i+1},\ i = 1,2,\ldots,n-1 \\ \frac{dx_n}{dt_0} &= \mu[f^-(X,w) + [f^+(X,w) - f^-(X,w)]\chi_1] \\ \frac{d\chi_j}{dt_0} &= \chi_{j+1},\ j = 1,2,\ldots,q-1 \\ \frac{d\chi_q}{dt_0} &= -k[f^+(X,w) - f^-(X,w)]\chi_1 - d_1\chi_2 \\ &\quad -\cdots - d_{q-1}\chi_q + k[F(X,r) - f^-(X,w)] \end{aligned} \tag{13}$$

results. By setting $\mu = 0$ we get the system given by

$$\begin{aligned} \frac{dx_i}{dt_0} &= 0,\ i = 1,2,\ldots,n \\ \frac{d\chi_j}{dt_0} &= \chi_{j+1},\ j = 1,2,\ldots,q-1 \\ \frac{d\chi_q}{dt_0} &= -k[f^+(X,w) - f^-(X,w)]\chi_1 - d_1\chi_2 \\ &\quad -\cdots - d_{q-1}\chi_q + k[F(X,r) - f^-(X,w)] \end{aligned} \tag{14}$$

Then the inverse replacement $t_0 = \mu^{-1}t$ yields the fast-motion subsystem (FMS) given by

$$\begin{aligned} \mu\dot{\chi}_j &= \chi_{j+1},\ j = 1,2,\ldots,q-1 \\ \mu\dot{\chi}_q &= -k[f^+(X,w) - f^-(X,w)]\chi_1 - d_1\chi_2 \\ &\quad -\cdots - d_{q-1}\chi_q + k[F(X,r) - f^-(X,w)] \end{aligned} \tag{15}$$

where X and w are treated as the frozen variables during the transients in (16). Finally, the equations (15) can be rewritten as

$$\begin{aligned} \mu^q \chi^{(q)} &+ d_{q-1}\mu^{q-1}\chi^{(q-1)} + \cdots + d_1\mu\chi^{(1)} \\ +k[f^+(X,w) &- f^-(X,w)]\chi = k[F(X,r) - f^-(X,w)]. \end{aligned} \tag{16}$$

Assumption 3: Assume that k, d_1,..., d_q, u^+, and u^- are selected such that the condition

$$k[f^+(X,w) - f^-(X,w)] > 0$$

holds for all $(X,w) \in \Omega_{X,w}$ and the transients of the FMS (16) are exponentially stable.

The rate of transients in (16) depends on the parameter μ. Hence, if μ is small enough, then, in accordance with Assumption 3, after the rapid decay of transients in (16), we get the steady state (more precisely, quasi-steady state) for the FMS, where $\dot{\chi} = 0$ and $\chi(t) = \chi^{id}(t)$. From (16), we find

$$\chi^{id} = \frac{F(X,r) - f^-(X,w)}{f^+(X,w) - f^-(X,w)} \tag{17}$$

where χ^{id} is exactly the inverse dynamics solution.

Substitution of $\chi = \chi^{id}$ into (7) yields the slow-motion subsystem (SMS) given by

$$x^{(n)} = F(X,r) \tag{18}$$

that is the reference model of the desired behavior in the following form:

$$\begin{aligned} T^n x^{(n)} &+ a_{n-1}^0 T^{n-1} x^{(n-1)} \\ &+ \cdots + a_2^0 T^2 x^{(2)} + a_1^0 T x^{(1)} + x = r. \end{aligned}$$

The main qualitative property of the singularly perturbed systems is that: if the equilibrium point of the FMS is exponentially stable, then there exists $\mu^* > 0$ such that for all $\mu \in (0, \mu^*)$ the trajectories of the singularly perturbed system approximate to the trajectories of the SMS [18], [19], [20], [21], [22], [23]. So, if a sufficient time-scale separation between the fast and slow modes in the closed-loop system and exponential convergence of FMS transients to equilibrium are provided, then after the damping of fast transients the desired output behavior prescribed by (18) is fulfilled despite that $f(X, w)$ and $g(X, w)$ are unknown complex functions. Thus, the output transient performance indices are insensitive to parameter variations of the nonlinear system and external disturbances, by that the solution of the discussed control problem (3) is maintained.

Let τ_{fms} be the time constant of the FMS (16) where $\tau_{fms} = \mu/\sqrt[q]{\gamma}$ and $\gamma = k[f^+(X, w) - f^-(X, w)]$. The requirement for degree of time-scale separation between the fast and slow modes in the system (12) can be represented by

$$\tau_{fms} \leq \tau_{sms}/\eta \tag{19}$$

where $\tau_{sms} = T$ and, for example, $\eta \geq 10$. The inequality (19) yields the upper bound for μ given by

$$\mu \leq \mu_{\max} = T\sqrt[q]{\gamma_{\min}}/\eta$$

where

$$\gamma_{min} = \min_{\forall (X,w) \in \Omega_{X,w}} k[f^+(X, w) - f^-(X, w)].$$

D. Control Variable Range

From (17) it follows, the not saturated behavior of $\chi^{id}(t)$ can be provided by proper selection of the pulse-width modulator parameters u^+ and u^-, it means that the value $|f^+(X, w) - f^-(X, w)|$ can be increased through the selection of u^+ and u^-.

In order to keep the condition $\chi \in [0, 1]$ for the transient behavior of the FMS (16), the control system should be supplemented by an additional limiter as shown in Fig. 2.

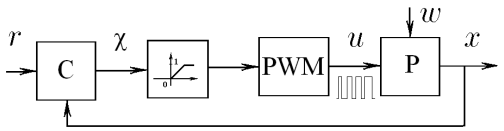


Fig. 2. Block diagram of control system which is supplemented by the limiter.

E. High-Frequency Chattering Attenuation

The high-frequency chattering, caused by the switching mode of the pulse-width modulator (2), is induced in the closed-loop system. The main effect is produced by the sampling frequency ω_s where $\omega_s = 2\pi/T_s$. The maximum impact of the pulse-width modulation on the amplitude of the high-frequency chattering with the sampling frequency ω_s in the behavior of the output variable $x(t)$ can be estimated by

$$A_x(\omega_s) \approx \frac{2}{\pi \omega_s^n} \mid f^+(X, w) - f^-(X, w) \mid \tag{20}$$

under assumption that $\chi = 0.5$.

Accordingly, the maximum impact of the pulse-width modulation on the amplitude of the high-frequency chattering with the sampling frequency ω_s in the behavior of the duty ratio function $\chi(t)$ can be estimated by

$$A_\chi(\omega_s) \approx \frac{k}{\mu^q \omega_s^{q-n}} A_x(\omega_s) \tag{21}$$

where $q \geq n$.

Note, the accuracy of the estimations given by (20) and (21) increase when $\omega_s \to \infty$.

From (20) and (21), we have $A_x(\omega_s) \to 0$ and $A_\chi(\omega_s) \to 0$ as $T_s \to 0$. Hence, the high-frequency chattering attenuation can be provided in the closed-loop system when $T_s \to 0$.

The effect of the generated high-frequency harmonics and subharmonic of unknown frequencies can be neglected in comparison with the effect of the sampling frequency ω_s due to a low-pass filtering property of the system (1).

Remark 6: From (21) it follows, the high-frequency chattering of the duty ratio function $\chi(t)$ can be reduced without decrease in T_s, by increasing q in comparison with n.

IV. Example: Control of Magnetic Levitation System

A. Model of Magnetic Levitation System

In this section the problem of tracking control for magnetic levitation system is discussed as an example of the application for the presented above PWM control design methodology. Let us consider a simplified model of the magnetic levitation system given by the following differential equations [25]:

$$\begin{aligned} \dot{x}_1 &= x_2 \\ \dot{x}_2 &= g_a - \frac{1}{M}\frac{x_3^2}{x_1} \\ \dot{x}_3 &= -\frac{R}{L}x_3 + \frac{1}{L}u \end{aligned} \tag{22}$$

where x_1 is the mass position relative to the electromagnet, x_2 is the mass velocity, x_3 is the electric current in the electromagnet coil, g_a is the gravity acceleration, M is the mass, R and L are the resistance and inductance of the electromagnet coil.

Assume that x_1 is the measured output variable and the scalar control variable $u(t)$ takes one of two possible values in $\Omega_u := \{u^-, u^+\}$ where the pulse-width modulated control for the magnetic levitation system (22) is defined as the switching function $u(t)$ given by (2). Accordingly, from (22) the Filippov's average model

$$\begin{aligned} \dot{x}_1 &= x_2 \\ \dot{x}_2 &= g_a - \frac{1}{M}\frac{x_3^2}{x_1} \\ \dot{x}_3 &= -\frac{R}{L}x_3 + \frac{1}{L}u^- + \frac{1}{L}[u^+ - u^-]\chi \end{aligned} \tag{23}$$

results. Denote

$$y_1 = x_1,\ y_2 = x_2,\ y_3 = g_a - \frac{1}{M}\frac{x_3^2}{x_1}. \tag{24}$$

Assume that the conditions $y_1 > 0$, $x_3 > 0$ hold, then

$$y_3 < g_a. \tag{25}$$

Hence, in accordance with (23)-(25), the Filippov's average model can be rewritten as

$$\begin{aligned} \dot{y}_1 &= y_2 \\ \dot{y}_2 &= y_3 \\ \dot{y}_3 &= f(y_1, y_2, y_3, u^-) + g(y_1, y_2, y_3, u^-, u^+)\chi \end{aligned} \tag{26}$$

where

$$f(y_1, y_2, y_3, u^-) = \left[\frac{y_2}{y_1} + \frac{2R}{L}\right](g_a - y_3) - \frac{2\sqrt{g_a - y_3}}{\sqrt{M}L\sqrt{y_1}}u^-$$

$$g(y_1, y_2, y_3, u^-, u^+) = \frac{(-1)2\sqrt{g_a - y_3}}{\sqrt{M}L\sqrt{y_1}}[u^+ - u^-]$$

B. Controller for Magnetic Levitation System

In accordance with (8), consider the controller given by

$$\begin{aligned} &\mu^3 \dddot{\chi} + d_2\mu^2 \ddot{\chi} + d_1\mu \dot{\chi} \\ &= k[-\dddot{y}_1 - \frac{a_2^0}{T}\ddot{y}_1 - \frac{a_1^0}{T^2}\dot{y}_1 + \frac{1}{T^3}(r - y_1)]. \end{aligned} \tag{27}$$

In order to practical implementation, the discussed control law (27) can be rewritten in the form given by

$$\begin{aligned} &\chi^{(3)} + a_2\chi^{(2)} + a_1\chi^{(1)} \\ &= b_3 y_1^{(3)} + b_2 y_1^{(2)} + b_1 y_1^{(1)} + b_0 y_1 + c_0 r \end{aligned} \tag{28}$$

where

$$a_2 = \frac{d_2}{\mu}, \quad a_1 = \frac{d_1}{\mu^2}, \quad c_0 = \frac{k}{\mu^3 T^3},$$

$$b_3 = -\frac{k}{\mu^3}, \quad b_2 = -\frac{ka_2^0}{\mu^3 T}, \quad b_1 = -\frac{ka_1^0}{\mu^3 T^2}, \quad b_0 = -c_0.$$

Then, from (28), we may get the equations of the controller in the state space form (10), that are

$$\begin{aligned} \dot{\chi}_1 &= -a_2\chi_1 + \chi_2 + (b_2 - a_2 b_3)y_1 \\ \dot{\chi}_2 &= -a_1\chi_1 + \chi_3 + (b_1 - a_1 b_3)y_1 \\ \dot{\chi}_3 &= -a_0\chi_1 + (b_0 - a_0 b_3)y_1 + c_0 r \\ \chi &= \chi_1 + b_3 y_1. \end{aligned} \tag{29}$$

C. Selection of Initial Conditions

In order to avoid excitation of transients caused by the initial condition mismatching between (22) and (29), let the initial conditions of the magnetic levitation system (22) and the controller given by (29) have been selected such that the steady-state of the closed-loop system composed of the Filippov's average model (23) and controller (29) takes place when $t = 0$. Hence, from (23) under assumptions $\dot{x}_1(0) = 0$, $\dot{x}_2(0) = 0$, $\dot{x}_3(0) = 0$, and $y_1(0) = x_1(0) = x_1^0 > 0$, it follows that $x_2(0) = x_3(0) = 0$ and

$$\begin{aligned} x_3(0) &= \sqrt{M g_a x_1(0)} \\ \chi(0) &= \frac{Rx_3(0) - u^-}{u^+ - u^-}. \end{aligned} \tag{30}$$

Then, from (29) under assumptions $\dot{\chi}_1(0) = 0$, $\dot{\chi}_2(0) = 0$, and $\dot{\chi}_3(0) = 0$, the following expressions

$$\begin{aligned} \chi_1(0) &= \chi(0) - b_3 y_1(0) \\ \chi_2(0) &= a_2\chi_1(0) - (b_2 - a_2 b_3)y_1(0) \\ \chi_3(0) &= a_1\chi_1(0) - (b_1 - a_1 b_3)y_1(0) \end{aligned} \tag{31}$$

result.

D. Simulation Results

The parameters of the discussed magnetic levitation system (22) are taken from [25] where we have

$$M = 0.1\,\text{kg}, \quad g_a = 9.8\,\text{m/s}^2, \quad R = 6\,\Omega, \quad L = 0.1\,\text{H}.$$

Assume that the pulse-width modulated control is defined by (29) with the switching function $u(t)$ given by (2) where

$$T_s = 0.006\,\text{s}, \quad u^+ = 2.2\,\text{V}, \quad u^- = 0.2\,\text{V}.$$

The controller parameters are selected as

$$\begin{aligned} &T = 0.1\,\text{s}, \quad a_1^0 = a_2^0 = 3, \quad d_1 = 15, \\ &d_2 = 8, \quad \mu = 0.0067\,\text{s}, \quad k = -0.015 \end{aligned}$$

The initial conditions for numerical simulation are selected in accordance with (30) and (31) where $x_1(0) = 0.07\,\text{m}$.

The simulation results of the closed-loop magnetic levitation control system (22), (29) with the switching function (2) are shown in Figs. 3–5. The simulation results confirm the presented above analytical investigations. In particular, the transient response of $x_1(t)$ coincided with the transient response of the reference equation assigned by

$$T^3\dddot{x}_1 + a_2^0 T^2 \ddot{x}_1 + a_1^0 T\dot{x}_1 + x_1 = r.$$

From (20) and (21), we have $A_x(\omega_s) \approx 10^{-6}$ and $A_\chi(\omega_s) \approx 0.0524$. One can see, the calculated value of $A_\chi(\omega_s)$ coincided with the high-frequency chattering of the duty ratio function $\chi(t)$ shown in Fig. 5.

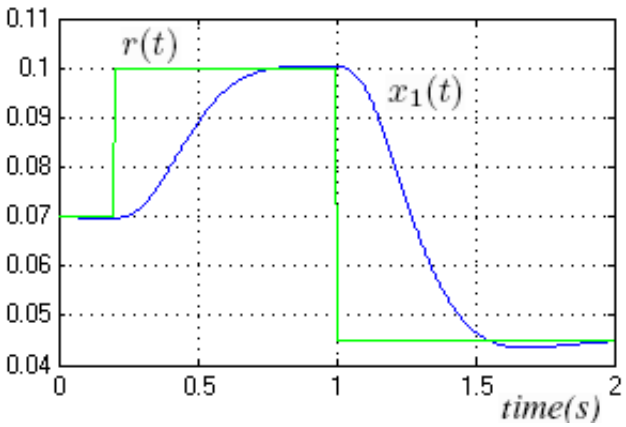


Fig. 3. Plots of $r(t)$ and $x_1(t)$ in the system (22), (29) with the switching function (2) where $t \in [0, 2]$ s.

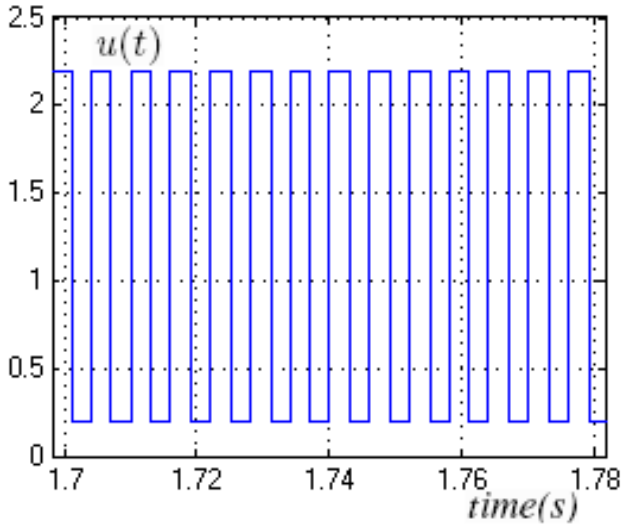


Fig. 4. Plot of $u(t)$ in the system (22), (29) with the switching function (2) where $t \in [1.7, 1.78]$ s.

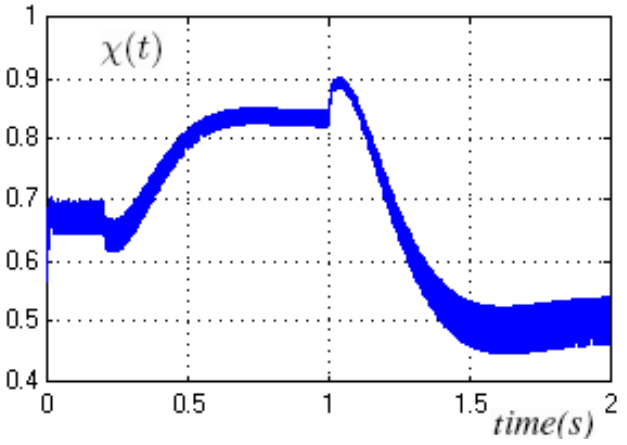


Fig. 5. Plots of $\chi(t)$ in the system (22), (29) with the switching function (2) where $t \in [0, 2]$ s.

V. Conclusion

The main advantage of the discussed singular perturbation technique for control system analysis and design is that the parameters of the controller for nonlinear systems in the presence of a pulse-width modulation in feedback loop can be analytically derived in accordance with such indirect performance objectives as the desired root placement of the reference model characteristic polynomial, while the desired root distribution is defined by such direct output performance objectives as settling time and overshoot. The application of the singular perturbation technique in the presented design methodology allows to get desired output transient performance indices for nonlinear systems under uncomplete knowledge about external disturbances and plant's parameter variations.

The next step of investigation will be an implementation of the discussed magnetic levitation control system in an experimental prototype in the laboratory.

Acknowledgment

This work was supported by Russian Foundation for Basic Research (RFBR) under grant no. 08-08-00982-a.

References

[1] J. Holtz, "Pulsewidth modilation: a survey," *IEEE Trans. Idustrial Electronics*, vol. IE-39, no. 5 pp. 410–420, 1992.

[2] P. Muir, and C. P. Neumann, "Pulsewidth modilation control of brushless DC motor for robotic applications," *IEEE Trans. Idustrial Electronics*, vol. IE-32, no. 3 pp. 222–229, 1985.

[3] L. Malesani, and P. Tomasin, "PWM current control techniques of voltage source converters - a survey," in *Proc. of the IECON'93 Int. Conf. on Idustrial Electronics, Control, and Instrumentation*, vol. 2, pp. 670–675, 1993.

[4] N. Mohan, T. M. Undeland, and W. P. Robbins, *Power Electronics: Converters, Applications and Design*, Wiley, 1995.

[5] R. E. Andeen, "The principle of equivalent areas," *Trans. of AIEE*, vol. 79, pp. 532–536, 1960.

[6] Y. Pan, Y. Uno, K. Furuta, and A. Ohata, "VSS-type self tuning PWM control," in *Proc. of American Control Conference*, 14-16 June, 2006.

[7] A. Kh. Gelig, and A. N. Churilov, "Pulse-Modulated Systems: a Review of Mathematical Approaches," *Functional-Differential Equations*, vol. 3, no. 3-4, pp. 267-320, 1996.

[8] A. Kh. Gelig, and A. N. Churilov, *Stability and Oscillations of Nonlinear Pulse-Modulated Systems*, Birkhauser, Boston, 1998.

[9] H. Sira-Ramirez, "A geometric approach to pulse-width-modulated control in nonlinear dynamical systems," *IEEE Trans. Automatic Control*, **34**, no. 2 pp. 184–187, 1989.

[10] H. Sira-Ramirez, and P. Lischinsky-Arenas, "Dynamical discontinuous feedback control of nonlinear systems," *IEEE Trans. Automatic Control*, **35**, no. 12 pp. 1373–1378, 1990.

[11] A. F. Filippov, "Differential equations with discontinuous right hand sides," *Am. Math. Soc. Transl.*, vol. 42, pp. 199–231, 1964.

[12] K. J. Åström and T. Hagglund, *PID controllers: Theory, Design, and Tuning*, Research Triangle Park, NC: Instrum. Soc. Amer., 1995.

[13] A. O'Dwyer, *Handbook of PI and PID Tuning Rules*, London, UK: Imperial College Press, 2003.

[14] G. Ziegel, and N. B. Nichols, "Optimum settings for automatic controllers," *Trans. ASME*, **64**, no. 8 pp. 759–768, 1942.

[15] *The Control Handbook*, Ed. by W.S. Lewine. Piscataway, NJ: CRC Press/IEEE Press, 1996.

[16] K. J. Åström, T. Hagglund, C. C. Hang, and W. K. Ho. "Automatic tuning and adaptation for PID controllers - A survey," *Contr. Eng. Pract.*, Vol. 1, no. 4, pp. 699–714, 1993.

[17] V. D. Yurkevich, *Design of Nonlinear Control Systems with the Highest Derivative in Feedback*, World Scientific, 2004.

[18] A. N. Tikhonov, "Systems of differential equations containing a small parameter multiplying the derivative," *Mathematical Sb.*, Moscow, vol. 31, no. 3, pp. 575–586, 1952.

[19] A. I. Klimushchev and N. N. Krasovskii, "Uniform asymtotic stability of systems of differential equations with a small parameter in the derivative terms," *J. Appl. Math. Mech.*, vol. 25, pp. 1011–1025, 1962.

[20] F. C. Hoppensteadt, "Singular perturbations on the infinite time interval," *Trans. of the American Mathematical Society*, vol. 123, pp. 521–535, 1966.

[21] P. V. Kokotović, "Applications of singular perturbation techniques to control problems," *SIAM Review*, vol. 26, no. 4, pp. 501–550, 1984.

[22] P. V. Kokotović, H. K. Khalil, and J. O'Reilly, *Singular Perturbation Methods in Control: Analysis and Design*, Philadelphia, PA: SIAM, 1999.

[23] D. S. Naidu, "Singular perturbations and time scales in control theory and applications: An overview," *Dynamics of Continuous, Discrete & Impulsive Systems (DCDIS), Series B: Applications & Algorithms*, vol. 9, no. 2, pp. 233–278, 2002.

[24] V. D. Yurkevich, "PWM PI/PID/PIDF Control for Nonlinear Nonaffine Systems via Singular Perturbation," *Proc. of the 2009 Int. Forum on Strategic Technologies (IFOST-2009)*, Ho Chi Minh city, Vietnam, 21-23 October, 2009, pp.185-190.

[25] M. R. Filho, and C. J. Munaro, "A design methodology of tracking controllers for magnetic levitation systems," in *Proc. of the 2001 IEEE Int. Conf. on Control Applications (CCA '01)*, 5-7 Sept., 2001, pp.47–51.

17 Robust Formation Controller Synthesis with Different Time Delays

Robust Formation Controller Synthesis with Different Time Delays

Ulf Pilz, Herbert Werner
Hamburg University of Technology
Institute of Control Systems
Eissendorfer Str. 40, D-21073 Hamburg, Germany
Email: {ulf.pilz, h.werner}@tu-harburg.de

***Abstract*— In this paper we consider the robust controller synthesis problem for a group of identical agents which have to fulfill a common goal. To achieve this, the agents have the capability to communicate with each other. The communication structure is modeled as a directed graph and is affected by time delays and changes in the communication topology. These time delays differ depending on whether they affect also an agent's own states and not only the communicated states of the other agents. For both cases, a synthesis method that guarantees stability of the formation is proposed. The design of a controller that guarantees stability in the face of self-delays is the main contribution of this paper and an extension to previous work. A comparison of the controller design method for communication delays with and without self-delays has been carried out by a simulation of a formation flight of quad-rotor helicopters.**

I. Introduction

The problem of developing cooperative control strategies for multi-agent systems consisting of identical agents has received considerable interest in the control community over the last several years due to its broad applications in mobile robotics, autonomous underwater vehicles, automated highway systems, microsatellite clustering, unmanned aerial vehicles, etc. [1], [2], [3], [4]. In all these applications it is important to design appropriate control strategies based on local information such that the overall goal of the multi-agent system can be fulfilled. To obtain the required information from other agents, communication has to be considered as a necessary part in real applications and has to be taken into account in the control strategy.

There are two important aspects of communication which shall be examined here. First, there is the effect of changing communication topologies due to range limitations, obstacles, packet losses or failures of the communication equipment. Another aspect is the delay introduced by communication over a certain channel. Since the aspect of changing communication topologies has already been covered in previous work [5], [6], in this paper we will concentrate on communication time delays. A lot of research effort has been directed to the study of communication delays. Multi-agent systems consisting of single integrator agents were considered in [7], [8], [9]. But also higher order agent dynamics were considered in [10], [11], [12], [13]. In contrast to these works and a previous result obtained in [6], in this paper the effect of two implementations of time delays shall be analysed and a local controller synthesis method taking these two different time delays into account is proposed. In the first case, the time delay only affects the states of other agents, such that the formation error signal will be composed of the difference of an agent's actual state and the delayed transmitted outputs of the other agents. However, in practice it is more realistic to consider the formation error as the difference between one agents' delayed outputs and the delayed outputs of the other agents which are delayed by the same amount. In both cases the time delays are allowed to be different. For both implementations the controller synthesis is carried out using methods from robust control.

This paper is organized as follows. In Section II a review of the multi-agent system framework is given and the different implementations of time delays is presented. In Section III the robust controller design approach for the two different implementations of time delay is given. Section IV presents a numerical example applied on a linear time-invariant model of a quad-rotor helicopter. Section V concludes this paper and addresses future work.

The following notation will be used throughout the paper: I_p denotes the $p \times p$ identity matrix; $\mathbb{R}^{p\times q}$, $\mathbb{C}^{p\times q}$ are, correspondingly, the sets of $p \times q$ real and complex matrices; $\mathrm{diag}\,(x_1, \ldots, x_n)$ indicates an $n \times n$ diagonal matrix with the elements x_1 to x_n on the diagonal; $\otimes$ is the Kronecker product; $\mathcal{F}_L(P(s), K(s))$ denotes the lower linear fractional transformation of $P(s)$ with $K(s)$, see, e.g., [14]. $\mathcal{D}_\tau$ represents the delay operator such that $\mathcal{D}_\tau v(t) = v\,(t-\tau)$ for a continuous function $v(t)$.

II. Stability of the Multi-Agent System and Formation Modeling

In this section a review of the proposed formation control framework is given and results from graph theory which are used to model the communication topology are breifly recalled. Furthermore we will extend the description of the communication topology by time delays.

A. Description of a Single Agent

We consider a formation of N identical agents which have the ability to communicate with each other. Each agent can be modeled as a LTI system and has a two-degree of freedom controller as shown in Fig. 1. The controller consists of two components:

- K_L is a feedback controller which is used to *internally* stabilize the agents, e.g., in case of loss of all communication.
- $K_F(s)$ is a controller which uses signals that are of interest on *formation level* and influences the formation behaviour.

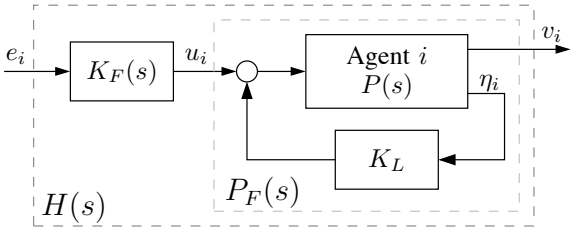


Figure 1. Single agent and its controller

Here, $v_i \in \mathbb{R}^p$, $i = 1, \ldots, N$, are the transmitted signals that are of interest for the formation of agents, whereas $\eta_i \in \mathbb{R}^m$ are outputs that can be used to stabilize the agent internally. Considering for example an unmanned aerial vehicle as will be done in Section IV, the transmitted signals may consist of the positions in cartesian coordinates of the vehicles whereas the internal singals may include additional information - such as angles, velocities or angular velocities. This also reduces the communication requirements that are needed to maintain a prescribed formation. The formation-level control error $e_i \in \mathbb{R}^p$ will be defined in the next subsection. The control signal $u_i \in \mathbb{R}^h$ is the input for agent i.

The transfer function from e_i to v_i is $H(s)$. The agents are defined to be locally stable if the eigenvalues of the system matrix of $H(s)$ are located strictly in the open left half plane $\mathcal{P} := \{\alpha | \alpha \in \mathbb{C}, \mathfrak{Re}\{\alpha\} < 0\}$.

B. Formation-Level Error and Communication Topology

For simplicity we first assume that there are no communication delays. The formation-level control error shall be defined as in [2]. It is the equally weighted sum of errors of the sensed neighbors. The set of sensed neighbors is given as $\mathcal{J}_i$ and consists of all agents k from which agent i receives information. $|\mathcal{J}_i|$ is the cardinality of this set.

$$e_i = \frac{1}{|\mathcal{J}_i|} \sum_{k \in \mathcal{J}_i} e_{ik}, \tag{1}$$

where

$$e_{ik} = (r_i - v_i) - (r_k - v_k) = \bar{r}_{ik} - (v_i - v_k) \tag{2}$$

is defined as the error between the i-th and k-th agent. The term $\bar{r}_{ik} \in \mathbb{R}^p$ defines the intended formation of the multi-agent system and consists of the difference between $r_i \in \mathbb{R}^p$ and $r_k \in \mathbb{R}^p$ which are the reference input for agent i and k, respectively. Note that the signals are weighted equally but the results in this paper can also be generalized if the weights in each communication channel are different as long as e_i is normalized.

The communication topology can be represented as a directed graph, where the nodes represent the agents and the vertices indicate communication links. To formally describe the communication topology one can use the normalized graph Laplacian matrix [15] which is defined as

$$L := [l_{ik}]_{N \times N}, \qquad l_{ik} = \begin{cases} 1, & \text{if } i = k \\ -\frac{1}{|\mathcal{J}_i|}, & k \in \mathcal{J}_i \\ 0, & k \notin \mathcal{J}_i. \end{cases} \tag{3}$$

C. Formation stability

Now we are able to construct the closed-loop formation shown in Fig. 2. Let $\mathbf{L_{(p)}} = L \otimes I_p$, $\mathbf{r} = \left[r_1^{\mathrm{T}} \ldots r_N^{\mathrm{T}}\right]^{\mathrm{T}}$, $\mathbf{e} = \left[e_1^{\mathrm{T}} \ldots e_N^{\mathrm{T}}\right]^{\mathrm{T}}$ and $\mathbf{v} = \left[v_1^{\mathrm{T}} \ldots v_N^{\mathrm{T}}\right]^{\mathrm{T}}$. Using the definition of the normalized Laplacian one can define $\mathbf{r}$ as a reference signal that provides a commanded value for the outputs of the agents whereas $\bar{\mathbf{r}}$ can be used as a reference signal for relative errors of the agents.

$$\mathbf{e} = \mathbf{L_{(p)}}\tilde{\mathbf{e}} = \mathbf{L_{(p)}}(\mathbf{r} - \mathbf{v}) = \bar{\mathbf{r}} - \mathbf{L_{(p)}}\mathbf{v}, \tag{4}$$

where $\tilde{\mathbf{e}}$ describes an absolute error of the outputs. Since the

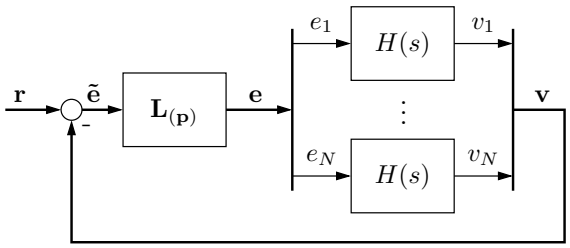


Figure 2. Closed-loop representation of the formation

normalized Laplacian L is not invertible, there are infinitely many absolute references $\mathbf{r}$ for the same relative reference $\bar{\mathbf{r}}$. For the closed-loop interconnection of the multi-agent system given in Fig. 2 we can define stability of the formation.

Definition 1 *A multi-agent system is called stable if the system matrix of the closed-loop transfer function from* $\mathbf{e}$ *to* $\mathbf{v}$ *has all its poles in the left half plane* $\mathcal{P}$.

The closed-loop system $P_F(s)$ formed by each agent and its stabilizing local controller K_L shall be described as

$$\begin{aligned} \dot{\xi}_i &= A\xi_i + Bu_i \\ v_i &= C\xi_i \qquad\qquad i = 1, \ldots, N \end{aligned} \tag{5}$$

where $A \in \mathbf{R}^{n \times n}$, $B \in \mathbf{R}^{n \times h}$ and $C \in \mathbf{R}^{p \times n}$. $\xi_i \in \mathbf{R}^n$ represent the states of an agent. With this definition the following result is taken from [2], Theorem 3.

Theorem 1 *The controller* K_F *stabilizes the closed-loop formation 2 if and only if it simultaneously stabilizes the set of* N *systems*

$$\begin{aligned} \dot{\tilde{\xi}}_i &= A\tilde{\xi}_i + B\tilde{u}_i \\ \tilde{v}_i &= \lambda_i C\tilde{\xi}_i \qquad\qquad i = 1, \ldots, N \end{aligned} \tag{6}$$

where λ_i denotes an eigenvalue of L. Note that due to the transformation (e.g., eigenvalue or Schur transformation, see [2] for details) of the system matrices the states, inputs and outputs are different from those in (5).

D. Time delay modeling

In this section the modeling of the communication time delays is presented. First we assume that the communication time delay between agent i and agent k is $\tau_{ik} \in \mathbb{R}_+$ for $i,k = 1,\ldots,N, i \neq k$. This means that the ouptut signal from each agent to itself is not delayed (agents without self-delay, [13]). The delay-dependent normalized graph Laplacian can be described as

$$L^d := \left[l_{ik}^d\right]_{N\times N}, \quad l_{ik}^d = \begin{cases} 1, & \text{if } i = k \\ -\frac{1}{|\mathcal{J}_i|}\mathrm{e}^{-s\tau_{ik}}, & k \in \mathcal{J}_i \\ 0, & k \notin \mathcal{J}_i. \end{cases} \tag{7}$$

With this definition the error between the i-th and k-th agent (2) changes to

$$e_{ik}^d(t) = \bar{r}_{ik}(t) - (v_i(t) - v_k(t-\tau_{ik})), \tag{8}$$

and the closed-loop formation can be shown to be as in Fig. 3, where $\mathbf{G_{(p)}^d} = (I_N \otimes I_p) - \mathbf{L_{(p)}^d}$ is the normalized adjacency matrix affected by communication time-delays and augmented with the number of outputs p.

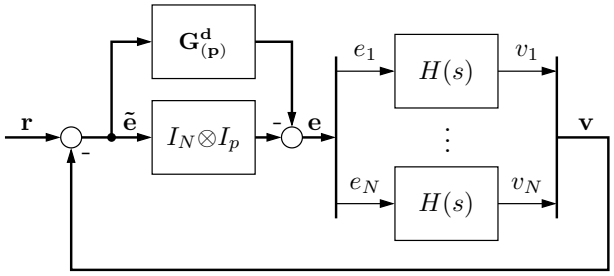


Figure 3. Formation with communication delays

The main contribution of this paper is an extention to the previous robust controller design approach given in [5], [6] to include agents with self-delays. This is a more realistic assumption in practice because one would not compare the actual position of one agent with the delayed postion of other agents which may lead to problems, e.g., concerning the avoidance of collisions between the agents. In this case the formation-level control error (2) becomes

$$e_{ik}^s(t) = \bar{r}_{ik}(t) - (v_i(t-\tau_{ik}) - v_k(t-\tau_{ik})). \tag{9}$$

Note that in both cases the relative reference input is not delayed since this would not change any of the obtained results.

III. Robust Controller Design

The robust controller design approach which is used here is proposed in [5] and [6]. In contrast to the methods proposed in [16] and [17] which consider only state feedback control, we rely on output feedback. Another distributed controller design method is proposed in [18] but this approach can be computationally expensive for large formations or agents with many states.

First we use a synthesis method already proposed in [5], [6] and incorporate communication time delays into the framework, see [19]. Then we extend this approach with another result from robust control theory [20] to stabilize formations including self-delays which is known to be a more challenging problem, see [13], [21].

A. Time-Varying Topologies and Communication Delays

In order to synthesize a robust controller we need to recall that the eigenvalues λ_i, $i = 1\ldots N$, of the normalized Laplacian L defined in (3) belong to the set [15]

$$\mathbf{\Lambda} := \{1 + \delta_\lambda | \delta_\lambda \in \mathbb{C}, |\delta_\lambda| \leq 1\}. \tag{10}$$

Proof: Communication delays without self-delays influence only the off-diagonal elements of L^d. In Fig. 3 the eigenvalues of the normalized adjacency matrix G are multiplied by a factor $\mathrm{e}^{-s\tau}$ which means that the magnitude of the eigenvalues do not change but only the phase according to the time-delay τ. Therefore all eigenvalues of L^d still belong to the set defined in (10). ■

Now we can apply a standard approach from robust control [14] already applied in [6], [19] and represent the group of N agents as a single one with uncertainty δ_λ.

Theorem 2 *A controller $K_F(s)$ stabilizes the closed-loop formation in Fig. 3 for any number of agents and any fixed communication topology if $K_F(s)$ stabilizes the following system $G^d(s)$ (see Fig. 4(a))*

$$\begin{aligned} \dot{\tilde{\xi}} &= A\tilde{\xi} + B\tilde{u} \\ z_\delta &= C_\delta \tilde{\xi} \\ \tilde{v} &= C\tilde{\xi} + D_\delta w_\delta \\ w_\delta &= \delta_\lambda I_p z_\delta \end{aligned} \tag{11}$$

for all $|\delta_\lambda| \leq 1$, where $D_\delta C_\delta = C$.

Consider now a norm bounded uncertainty $\Delta \in \mathbf{\Delta}$

$$\mathbf{\Delta} := \{\Delta \,|\, \Delta \in \mathbf{C}^{q\times q}, \|\Delta\| \leq 1\}$$

acting on $G^d(s)$, $w_\delta = \Delta z_\delta$ and accounting for the structure of the uncertainty (i.e., $\Delta = \delta_\lambda I_p$) the following theorem enables the controller design using μ-synthesis with static D-scaling.

Theorem 3 *A controller $K_F(s)$ stabilizes the closed-loop formation in Fig. 3 for any number of agents and any fixed as well as time-varying communication topology with any time-varying communication delays if there exists an invertible matrix $D^d \in \mathbf{R}^{p\times p}$ such that*

$$\|D^d T^d(s) D^{d^{-1}}\|_\infty < 1,$$

where $T^d(s) = \mathcal{F}_L\left(G^d(s), -K_F(s)\right)$ as shown in Fig. 4(a).

The lengthy proof of this theorem is presented in [19], Theorem 3 and is omitted here due to space limitations.

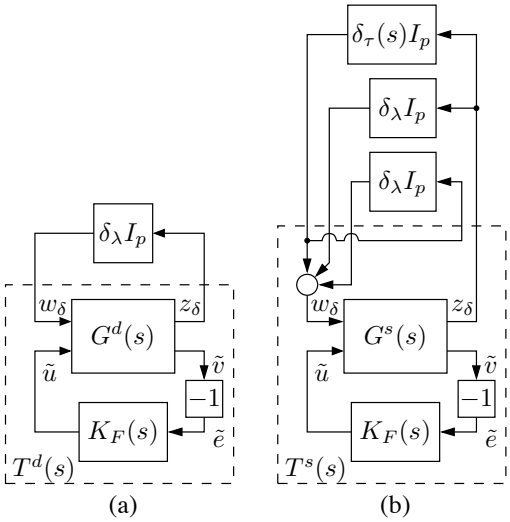


Figure 4. LFT interconnection for robust stability design

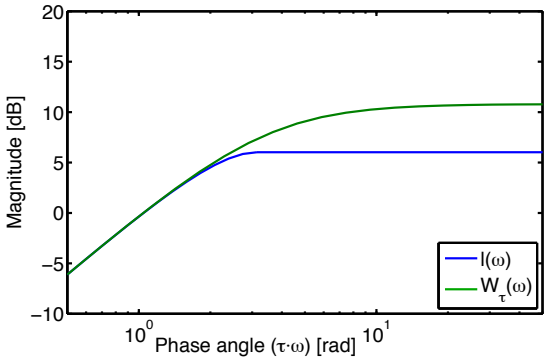


Figure 5. Multiplicative weight for time delay uncertainty τ

B. Incorporating Self-Delays

In this section we show the main result of this paper: how to incorporate self-delays in the presented framework and design a controller that guarantees stability of the formation. In contrast to Section III-A we use (9) to incorporate self-delays into the framework. Here we assume that the time delays are uncertain and may vary within a given range to be upper-bounded and this upper bound is considered to be known. Furthermore, we assume that the difference between one agent's own state and the state of an neighboring agent are delayed by the same amount. In practice this can be achieved if the communicated outputs of the agents include a time-stamp. Then the time delays can be modelled as an output complex multiplicative uncertainty where the tightest bound $l(\omega)$ of the multiplicative uncertainty weight can be described as [22]

$$l(\omega) = \begin{cases} \left|\mathrm{e}^{-(\mathrm{j}\omega)\tau} - 1\right|, & \forall \omega < \pi/\tau \\ 2, & \forall \omega \geqslant \pi/\tau. \end{cases} \tag{12}$$

For controller synthesis purposes we need to find a rational weight $W_\tau(s)$ that gives an upper bound for $l(\omega)$, i.e. $\left|\mathrm{e}^{-(\mathrm{j}\omega)\tau} - 1\right| < |W_\tau(s)|$. One easy way is to model the time delay as a first order weight $W_\tau(s)$ given in Appendix A. In Fig. 5 the irrational bound for the time delay and the first-order approximation is shown. One can see that the first order weight is very close to the irrational bound at low frequencies. At high frequencies there is a bigger difference but it has already been shown in [22] that $W_\tau(s)$ is a simple and reasonably good approximation of the time delay. The next theorem is adapted from Theorem 2 and a direct result of including the communication time delays with self-delays in the output channels $\tilde{v}$. Then the output equation in (6) changes to

$$\begin{aligned} \tilde{v}_i &= \lambda_i \mathcal{D}_{\tau_{ik}} I_p C \tilde{\xi}_i \\ &= \mathcal{D}_{\tau_{ik}} I_p C \xi_i + \delta_i \mathcal{D}_{\tau_{ik}} I_p C \xi_i. \end{aligned} \tag{13}$$

Following the same resoning as in Theorem 2 we are now ready to describe the group of N agents as a single one, now using δ_λ as an uncertainty describing the unknown communication topology and $\delta_\tau(s)$ as an uncertainty which corresponds to the unknown but upper bounded time delay τ.

Theorem 4 *A controller $K_F(s)$ stabilizes the closed-loop formation in Fig. 3 and including self-delays for any number of agents, any fixed communication topology and any fixed time delay τ if $K_F(s)$ stabilizes the following system $G^s(s)$ (see Fig. 4(b))*

$$\begin{aligned} \dot{\tilde{\xi}} &= A\tilde{\xi} + B\tilde{u} \\ z_\delta &= C_\delta \tilde{\xi} \\ \tilde{v} &= C\tilde{\xi} + D_\delta w_\delta \\ w_\delta &= \delta_\tau(s) I_p z_\delta + \delta_\lambda I_p z_\delta + \delta_\lambda I_p \delta_\tau(s) I_p z_\delta \end{aligned} \tag{14}$$

for all $|\delta_\lambda| \leq 1, \|\delta_\tau(s)\|_\infty \leq 1$ where $D_\delta C_\delta = C$.

Now using the same line of argument as in Theorem 3 and the results given in [20], the following theorem holds

Theorem 5 *A controller $K_F(s)$ stabilizes the closed-loop formation in Fig. 3 for any number of agents and any fixed as well as time-varying communication topology with any time-varying communication delays if there exists an invertible matrix $D^s \in \mathbf{R}^{3p\times 3p}$ such that*

$$\|D^s T^s(s) D^{s^{-1}}\|_\infty < 1,$$

where $T^s(s) = \mathcal{F}_L\left(G^s(s), -K_F(s)\right)$ as shown in Fig. 4(b).

This theorem leads to a scaled $\mathcal{H}_\infty$ condition and can be used for controller synthesis using standard robust control tools.

C. Performance Requirements

As has been shown in [6], stability by itself does not lead to a satisfactory control scheme. To have a meaningful formation-level controller, performance requirements have to be incorporated into the design. In this paper we will use a mixed-sensitivity approach. The generalized plant $\overline{G}(s)$ augmented with exogeneous inputs w_P and performance outputs z_P is shown in Fig. 6. In this figure, $\widetilde{P}_F(s)$ describes a state

space model of a system given in Theorem 1. To construct the generalized plant, one has to augment $G^d(s)$ or $G^s(s)$ with sensitivity $W_S(s)$ and control sensitivity $W_K(s)$ filters. Here we consider the exogeneous input w_P as a reference signal r. Note the fictitious input w_δ which accounts for the time delay uncertainty and the time-varying communication topology. To cope with unknown self-delays, we also introduce the filter $W_\tau(s)$ to augment the system $G^s(s)$ (see Fig. 6).

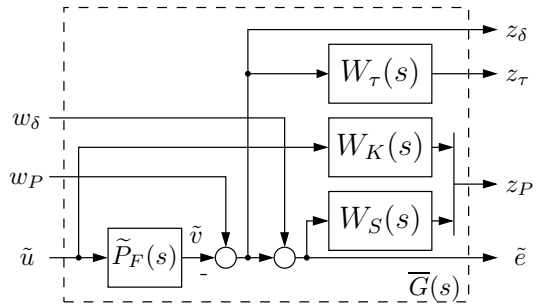


Figure 6. Generalized plant with additional channels

In this configuration the generalized plant $\overline{G}(s)$ imposes a penalty on the relative error $\tilde{e}$ of the agent in the formation but this can also be modified to establish a penalty on the absolute error.

IV. Numerical Example

This section shows an application of the proposed design method in Section III. We will both show the robustness against time delay uncertainties as well as changes in the communication topology. The first simulation will include time delays in the communication channel without self-delays. Unknown but upper-bounded self-delays will be added in the second simulation example

As an example of a multi-agent system we use a formation of quad-rotor helicopters. The number of agents in our simulation is $N = 5$. The linearized model of the quad-rotor helicopters has already been introduced in [6] together with the stabilizing local state feedback LQR controller K_L guaranteeing stability of one quad-rotor in case of a total communication loss. The formation controller is synthesized using μ-synthesis with static D-scaling. The sensitivity and control sensitivity filters and the filter for imposing an upper bound on the time delay uncertainty are given in Appendix A. In all simulations the communication topology changes randomly three times at 25, 50 and 75s. The communication time delay is assumed to be varying between 0 and 1.5s and it is assumed that the time delays for every agent are different. To force the quad-rotors to approach an absolute position we introduce a target waypoint with no dynamics which is known to at least one of the quad-rotors. In both simulations we restrict the figures to show the results only for the x-axis because there are no remarkable differences compared to the other axes.

The first simulation uses the controller synthesis approach suggested in Section III-A which leads to a 19th order controller. Fig. 7 shows the results of the simulation. Although there are some oscillations at the beginning of the simulation, the formation remains stable and also reaches the desired waypoint with a small steady state error. If we introduce even moderate self-delays (e.g., 0.75s) according to (9) into the closed-loop formation of quad-rotor helicopters it becomes unstable.

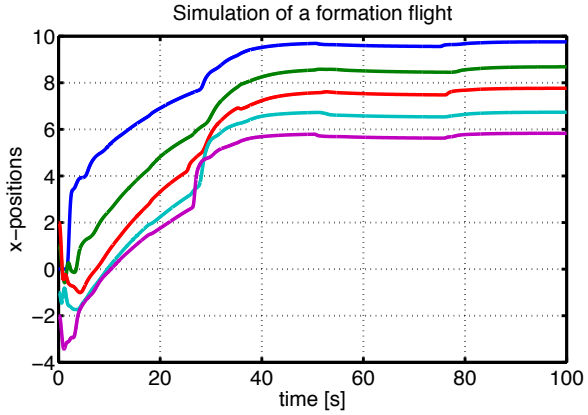


Figure 7. Simulated x-Positions of the formation without self-delays using synthesis method described in Section III-A

To cope with self-delays in the next simulation we apply the controller synthesis method proposed in Section III-B which leads to a 22nd order controller due to the introduction of the first order filter $W_\tau(s)$ as an upper bound for the maximum time-delay τ. Fig. 8 shows the simulation results of the formation flight example. In this figure one can see that the controller is able to stabilize the formation despite changes in the communication topology and communication time delays. The oscillations which were also present in Fig. 7 can also be observed in this simulation. In addition also the performance of this controller is slightly worse compared to the simulation where self-delays were not considered. This can also be explained intuitively because adding time delays tends to destabilize and degrade the performance of a system.

V. Conclusions and Future Work

This paper is concerned with the design of robust controllers for a formation of identical vehicles subject to communication and self-delays. The main contribution is the local controller synthesis for agents with self-delays that guarantee robustness against unknown but upper-bounded delays. Our design method is confirmed in a case study - in a numerical example a multi-agent system consisting of quad-rotor helicopters shall perform a formation flight.

In this framework the agents itself are stabilized internally by a local controller. Stability of the whole formation is guaranteed by local formation controllers that take into account time-varying communication topologies and arbitrary communication time delays. If the communication time delays

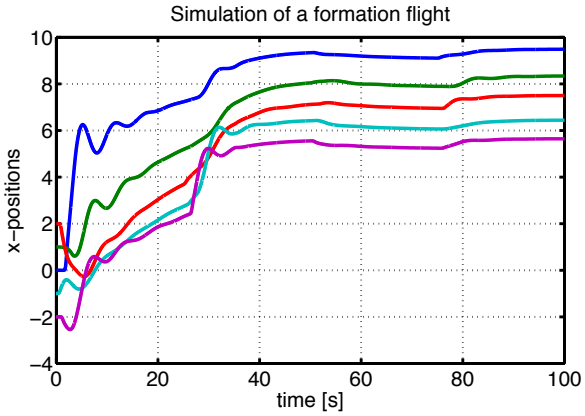


Figure 8. Simulated x-Positions of the formation including self-delays using synthesis method described in Section III-B

are assumed to be upper-bounded and this bound is known, the formation controller can also handle self-delays. Performance requirements are also incorporated into this design using mixed-sensitivity loop shaping.

Future research will concentrate on enhancing the performance of the robust controller design. This may be done by incorporating a prediction scheme. For this purpose the predictor then has to be robustified against uncertainties in the communication time delays.

APPENDIX

The sensitivity and control sensitivity weighting filters are given as

$$W_S = I_3 \otimes \left(\frac{0.333}{s+0.01}\right), \quad W_K = I_4 \otimes \left(10\frac{s+10^3}{s+10^6}\right).$$

The filter that describes the upper bound for the time-delay uncertainty is given as (see Fig. 5)

$$W_\tau = I_3 \otimes \left(\frac{3.465\tau s}{\tau s + 3.465}\right).$$

REFERENCES

[1] A. Jadbabaie, J. Lin, and S. Morse, "Coordination of groups of mobile autonomous agents using nearest neighbor rules," *IEEE Transactions on Automatic Control*, vol. 6, no. 48, pp. 988–1001, 2003.

[2] J. A. Fax and R. Murray, "Information flow and cooperative control of vehicle formations," *IEEE Transactions on Automatic Control*, vol. 49, pp. 1465–1476, 2004.

[3] W. Ren, R. Beard, and E. Atkins, "Information consensus in multivehicle cooperative control," *IEEE Control Systems Magazine*, vol. 27, no. 2, pp. 71–82, 2007.

[4] J. Wang and N. Elia, "Controller synthesis for reusable agents of consensus networks," in *Proc. 1st IFAC Workshop on Estimation and Control of Networked Systems*, Venice, Italy, 2009, pp. 102–107.

[5] A. Popov and H. Werner, "A robust control approach to formation control," in *Proc. European Control Conference*, Budapest, Hungary, 2009, pp. 4428–4433.

[6] U. Pilz, A. Popov, and H. Werner, "Robust controller design for formation flight of quad-rotor helicopters," in *Proc. 48th IEEE Conference on Decision and Control*, Shanghai, China, 2009, pp. 8322–8327.

[7] A. Papachristodoulou and A. Jadbabaie, "Synchronization of oscillator networks with heterogeneous delays, switching topologies and nonlinear dynamics," in *45th IEEE Conference on Decision and Control*, 2006, pp. 4307–4312.

[8] R. Ghabcheloo, A. Aguiar, A. Pascoal, and C. Silvestre, "Synchronization in multi-agent systems with switching topologies and non-homogeneous communication delays," in *46th IEEE Conference on Decision and Control*, 2007, pp. 2327–2332.

[9] P. Bliman and G. Ferrari-Trecate, "Average consensus problems in networks of agents with delayed communications," *Automatica*, vol. 8, no. 44, pp. 1985–1995, 2008.

[10] D. Lee and M. Spong, "Agreement with non-uniform information delays," in *Proc. American Control Conference*, Minneapolis, MN, USA, 2006, pp. 756–761.

[11] N. Chopra and M. Spong, *Advances in Robot Control*. Springer Verlag Berlin, 2006, ch. Passivity-based Control of Multi-agent Systems, pp. 107–134.

[12] W. Yang, A. Bertozzi, and X. Wang, "Stability of a second order consensus algorithm with time delay," in *47th IEEE Conference on Decision and Control*, 2008, pp. 2926–2931.

[13] U. Münz, A. Papachristodoulou, and F. Allgöwer, "Generalized Nyquist consensus condition for linear multi-agent systems with heterogeneous delays," in *Proc. 1st IFAC Workshop on Estimation and Control of Networked Systems*, Venice, Italy, 2009.

[14] K. Zhou and J. Doyle, *Essentials of Robust Control*. Upper Saddle River, NJ: Prentice-Hall, 1998.

[15] R. Horn and C. Johnson, *Matrix Analysis*. Cambridge: Cambridge University press, 1985.

[16] D. Gu, "A differential game approach to formation control," *IEEE Transactions on Control Systems Technology*, vol. 16, no. 1, pp. 85–93, 2008.

[17] F. Borelli and T. Keviczky, "Distributed LQR design for dynamically decoupled systems," in *Proc. 45th IEEE Conference on Decision and Control*, San Diego, CA, USA, 2006, pp. 5639–5644.

[18] P. Massioni and M. Verhaegen, "Distributed control for identical dynamically coupled systems: A decomposition approach," *IEEE Transactions on Automatic Control*, vol. 54, no. 1, pp. 124–135, 2009.

[19] A. Popov and H. Werner, "A necessary and sufficient condition for stability of a multi-agent system under any time-varying communication topology and time-varying delays," *IEEE Transactions on Automatic Control*, 2010, online: `http://www.tu-harburg.de/~rtsap/PoWe10TAC.html`.

[20] Y. Huang and K. Zhou, "Robust control of uncertain time delay systems," in *American Control Conference 1999*, 1999, pp. 1130–1135.

[21] Y. Tian and C. Liu, "Consensus of multi-agent systems with diverse input and communication delays," *IEEE Transactions on Automatic Control*, vol. 9, no. 53, pp. 2122–2128, 2008.

[22] Z. Wang, P. Lundström, and S. Skogestad, "Representation of uncertain time delays in the $\mathcal{H}_\infty$ framework," *International Journal of Control*, vol. 59, no. 3, pp. 627–638, 1994.

18 On adaptive control problems of continuous-time stochastic systems

On adaptive control problems of continuous-time stochastic systems

Uwe Küchler
Institute of Mathematics,
Humboldt University Berlin
Unter den Linden 6, D-10099, Berlin, Germany
Email: kuechler@math.hu-berlin.de

Vyacheslav A. Vasiliev
Department of Applied Mathematics and Cybernetics,
Tomsk State University,
Lenina 36, 634050 Tomsk, Russia
Email: vas@mail.tsu.ru

***Abstract*—Let $x = (x_t)_{t\geq 0}$ be a scalar observed process with control $u = (u_t)_{t\geq 0}$ described by the stochastic differential equation**

$$dx_t = \vartheta x_t dt + u_t dt + dw_t, \quad t \geq 0,$$

driven by the standard Wiener process $(w_t)_{t\geq 0}$. Assume the parameter ϑ to be unknown.

The problem solved in this paper is to approximate the process x to the stable Ornstein-Uhlenbeck process $x^0 = (x^0_t)_{t\geq 0}$ with the given dynamic parameter $a < 0$, satisfying the equation

$$dx^0_t = ax^0_t dt + dw_t, \quad t \geq 0,$$

by choosing the control process u.

More precisely, based on continuous observation of x, for given $\varepsilon > 0$ an adaptive control law $u^\varepsilon = (u^\varepsilon_t)_{t\geq 0}$ is constructed, such that the corresponding observed process $x^\varepsilon = (x^\varepsilon_t)_{t\geq 0}$, $x^\varepsilon_t = x_t(u^\varepsilon)$ satisfies the following relations

$$\sup_{t\geq 0} E(x^\varepsilon_t)^2 \leq L$$

and

$$\sup_{t\geq t_\varepsilon} E(x^\varepsilon_t - x^0_t)^2 \leq \varepsilon,$$

where L is some constant (independent of ε) and t_ε is an unboundedly increasing as $\varepsilon \to 0$ non-random function.

The first relation ensures the stability of the object x^ε.

The rate $\varepsilon^{-1}\ln\varepsilon^{-1}$ of increase of t_ε is obtained.

A similar problem is solved for a stochastic delay differential equation with an unknown parameter.

Keywords: Adaptive control; continuous-time stochastic systems; delay equations; unknown parameters

I. Introduction and problem statement

This paper is devoted to control problems of continuous-time stochastic processes with unknown parameters. A significant progress has been achieved in constructing the regulators with feedback for stochastic discrete and continuous time systems with unknown parameters, see e.g. [1],[2],[4]-[8], [11]-[20] among others. The adaptation to unknown parameters by using control leads to rather complicate analytical questions. Therefore it is quite natural to solve first the problem of estimating the unknown parameters of the object and then to use these estimators step by step in the algorithm of adaptive control.

In many cases the control aim can be given as 'a target inequality'

$$Q(x,t) \leq M, \tag{1}$$

where $Q(\cdot,\cdot)$ is a destination function, $x = (x_t)$ is a controlled process, M - a (given) threshold quantity, see [7], [8]. It is obvious, that for uncertainty systems (e.g. for the systems with unknown parameters) this inequality can be not fulfilled for all t. In such case the target inequality (1) can be changed by 'the limit target inequality'

$$\overline{\lim_{t\to\infty}}\, Q(x,t) \leq M. \tag{2}$$

Thus the control problem of an uncertain object can be formulated as follows: we have to find a control law which does not depend on unknown parameters such that the limit target inequality (1) or (2) is fulfilled (similar problems have been considered, e.g. in [7] Chap. 6, [8] Chap. 6, [19] Chap. 12, see the references therein as well).

This paper presents an adaptive control method for the following two problems.

Problem I.

Consider the stochastic differential equation with control given by

$$dx_t = \vartheta x_t dt + u_t dt + dw_t, \quad t \geq 0. \tag{3}$$

Here $(w_t)_{t\geq 0}$ denotes a realvalued standard Wiener process on some probability space $(\Omega, \mathcal{F}, P)$ with respect to a filtration $\underline{\mathcal{F}} = (\mathcal{F}_t,\ t \geq 0)$ from $\mathcal{F}$. We shall suppose in the sequel, that the initial value x_0 is zero mean Gaussian and $\mathcal{F}_0$-adapted. The control function $(u_t)_{t\geq 0}$ is supposed to be $(\mathcal{F}_t)$-adapted and satisfying for some constant K the inequalities $|u_t| \leq K\cdot|x_t|,\ t \geq 0$ (this condition ensures the existence of a strong solution of the equation (3)).

The parameter ϑ is supposed to be unknown but from the interval $[\vartheta_0, \vartheta_1]$, where $\vartheta_0,\ \vartheta_1$ are known numbers. Obviously, the process x can be unstable.

Consider the problem of approximation of the object $x = (x_t)_{t\geq 0}$ to the stable Ornstein-Uhlenbeck process $x^0 = (x^0_t)_{t\geq 0}$ (called a reference process) with a given dynamic parameter $a < 0$ satisfying the equation

$$dx^0_t = ax^0_t dt + dw_t, \quad x^0_0 = x_0, \quad t \geq 0. \tag{4}$$

It is clear, that in the case of known parameter ϑ the control law of the form

$$u_t = (a - \vartheta)x_t, \quad t \geq 0 \tag{5}$$

transform the equation (3) to (4) exactly.

It should be noted, that the linear structure (with respect to x) of the control function is usually used for quadratic type criteria. We shall use the quadratic cost criteria of the control performance as well, see formulae (7) and (8) below. Thus it is natural to use the function (5) for the case of unknown parameter ϑ with some non-anticipative estimator $\hat{\vartheta}_t$ instead of ϑ :

$$u_t = (a - \hat{\vartheta}_t)x_t, \quad t \geq 0. \tag{6}$$

Our main purpose is to obtain for a given positive ε a control law $u^\varepsilon = (u_t^\varepsilon)_{t\geq 0}$ from the admissible set of controls $u = u^\varepsilon$ of the type (6) with a $\mathcal{F}_t$-adapted dependent of ε process $(\hat{\vartheta}_t^\varepsilon)_{t\geq 0}$, $\hat{\vartheta}_t^\varepsilon \in [\vartheta_0, \vartheta_1]$, such that the corresponding observed process $x^\varepsilon = (x_t^\varepsilon)_{t\geq 0}$, $x_t^\varepsilon = x_t(u^\varepsilon)$ is stable in the following sense

$$\sup_{t\geq 0} E(x_t^\varepsilon)^2 \leq L, \tag{7}$$

where L is some constant (independent of ε), and approximate the reference process x^0 in the following way

$$\sup_{t\geq t_\varepsilon} E(x_t^\varepsilon - x_t^0)^2 \leq \varepsilon. \tag{8}$$

Here t_ε is an unboundedly increasing as $\varepsilon \to 0$ non-random function with a known rate of increase if ε decreases to zero.

Problem II.

Consider the stochastic delay differential equation with control given by

$$dx_t = \vartheta x_{t-r}dt + u_t dt + dw_t, \quad t \geq 0. \tag{9}$$

We shall suppose in the sequel, that the parameters ϑ, and $r > 0$ are real numbers; the initial process $x(0) = (x_0(s),\ s \in [-r, 0])$ also defined on (Ω, F, P) is cadlag and all $x_0(s),\ s \in [-r, 0]$ are zero mean Gaussian and $\mathcal{F}_0-$adapted.

The parameter ϑ is supposed to be unknown but such that $\vartheta \in [\vartheta_0, \vartheta_1]$, where $\vartheta_0,\ \vartheta_1$ are known bounds. The parameter r is known.

Consider the problem of approximation of the object (9) to the stable process $x^0 = (x_t^0)_{t\geq 0}$, satisfying the following stochastic delay differential equation with given dynamic parameters a and b :

$$dx_t^0 = ax_t^0 dt + bx_{t-r}^0 dt + dw_t, \quad t \geq 0 \tag{10}$$

and with the initial process $x(0)$.

Similar to Problem I we will use the obvious structure of the non-adaptive control law of the form

$$u_t = ax_t + (b - \vartheta)x_{t-r}, \quad t \geq 0$$

for the construction of an adaptive one:

$$u_t^\varepsilon = a_t^\varepsilon x_t + (b - \hat{\vartheta}_t^\varepsilon)x_{t-r}, \quad t \geq 0, \tag{11}$$

where bounded $\mathcal{F}_t$-adapted functions a_t^ε and $\hat{\vartheta}_t^\varepsilon \in [\vartheta_0, \vartheta_1]$, $t \geq 0$ are such that the corresponding observed process x^ε should satisfy the relations (7) and (8) with an unboundedly increasing as $\varepsilon \to 0$ non-random function t_ε.

The method used for solving both problems can be roughly described as follows.

First we shall construct an increasing sequence of stopping times and construct corresponding sequential maximum likelihood estimators of the parameter ϑ calculated at these times.

Secondly we define, using these estimators, a piecewise constant function $\vartheta^\varepsilon = (\vartheta_t^\varepsilon)_{t\geq 0}$ (and (a_t^ε) for Problem II) and a corresponding control function u^ε of the type (6) (and (11)).

It should be noted that an analogue method of parameter estimation was used in adaptive control and non-parametric estimation problems for discrete-time systems, see, for example, [21]-[24].

II. Main results

A. Construction of the control law for Problem I

We shall define a control law satisfying (7) and (8) as follows.

Let $(c_{n,\varepsilon})_{n\geq 0}$ be unboundedly increasing sequences of positive numbers, satisfying some special conditions (an example see after Corollary 2.1) and $(\tau_{n,\varepsilon})_{n\geq 0}$ be the sequences of stopping times

$$\tau_{n,\varepsilon} = \inf\{T > 0 : \int_0^T (x_t^\varepsilon)^2 dt = c_{n,\varepsilon}\}, \quad n \geq 0.$$

Define a sequence of sequential estimators $(\tilde{\vartheta}_{n,\varepsilon})_{n\geq 0}$ of the parameter ϑ as

$$\tilde{\vartheta}_{n,\varepsilon} = \frac{1}{c_{n,\varepsilon}} \int_0^{\tau_{n,\varepsilon}} x_t^\varepsilon (dx_t^\varepsilon - u_t^\varepsilon dt), \quad n \geq 0$$

and the piecewise-constant function $(\vartheta_t^\varepsilon)_{t\geq 0}$:

$$\vartheta_t^\varepsilon = \begin{cases} \vartheta_1, & 0 \leq t < \tau_{0,\varepsilon}, \\ \vartheta_{n,\varepsilon}, & \tau_{n,\varepsilon} \leq t < \tau_{n+1,\varepsilon}, \quad n \geq 0 \end{cases}$$

with

$$\vartheta_{n,\varepsilon} = (\tilde{\vartheta}_{n,\varepsilon} \wedge \vartheta_1) \vee \vartheta_0, \quad n \geq 0.$$

Here and in the sequel $a \wedge b$ =min(a,b), $a \vee b$ =max(a,b).

From the definition of the sequential plans $(\tau_{n,\varepsilon}, \vartheta_{n,\varepsilon}),\ n \geq 0$ of estimation of the parameter ϑ it follows that all the stopping times $\tau_{n,\varepsilon}$ are almost surely finite and according to the Burkholder-Gundy inequality with the smallest coefficient $\overline{C}(\theta) = (8\theta)^\theta$ (see, e.g. [3], [26]), for every $\theta \geq 1$ and $n \geq 0$ we have

$$E|\vartheta_{n,\varepsilon} - \vartheta|^{2\theta} \leq c_{n,\varepsilon}^{-2\theta} E \left(\int_0^{\tau_{n,\varepsilon}} x_t^\varepsilon dw_t \right)^{2\theta} \leq \overline{C}(\theta) c_{n,\varepsilon}^{-\theta}. \tag{12}$$

As the control law u^ε we define

$$u_t^\varepsilon = (a - \vartheta_t^\varepsilon) \cdot x_t^\varepsilon, \quad t \geq 0. \tag{13}$$

Our aim is to show that inserting u_t^ε from (13) into (3) the inequalities (7) and (8) are ensured.

To this end we decompose $x_t^\varepsilon - x_t^0$ into two parts:

$$x_t^\varepsilon - x_t^0 = \Delta_t^\varepsilon + \delta_t^\varepsilon, \quad t \geq 0 \tag{14}$$

and derive upper bounds for the second moments of Δ_t^ε and δ_t^ε. Here

$$\Delta_t^\varepsilon = y_t^\varepsilon - x_t^0, \quad \Delta_0^\varepsilon = 0,$$

$$\delta_t^\varepsilon = x_t^\varepsilon - y_t^\varepsilon, \quad \delta_0^\varepsilon = 0, \quad t \geq 0$$

and $y^\varepsilon = (y_t^\varepsilon)_{t\geq 0}$ is given by the SDE

$$dy_t^\varepsilon = A_t^\varepsilon y_t^\varepsilon dt + dw_t, \quad y_0^\varepsilon = x_0.$$

The auxiliary process $(A_t^\varepsilon)_{t\geq 0}$ is defined in a somewhat complicate way to ensure the desired estimators for (Δ_t^ε) and (δ_t^ε).

First we choose a piecewise-constant function $(\gamma_t^\varepsilon)_{t\geq 0}$ as follows:

$$\gamma_t^\varepsilon = \begin{cases} \vartheta_1 - \vartheta_0, & 0 \leq t < \tau_{0,\varepsilon}, \\ \varepsilon^{1/2}\gamma_{n,\varepsilon}, & \tau_{n,\varepsilon} \leq t < \tau_{n+1,\varepsilon}, \quad n \geq 0, \end{cases} \tag{15}$$

where $(\gamma_{n,\varepsilon})_{n\geq 0}$ is a decreasing sequence of non-random positive numbers, $\varepsilon > 0$.

Without restriction of generality we choose a number ε_0^* such that for all $\varepsilon \leq \varepsilon_0^*$ the following inequalities hold

$$\varepsilon^{1/2}\gamma_{0,\varepsilon} \leq \pi(\varepsilon) < -a, \tag{16}$$

where $\pi(\varepsilon) = (\ln(\mathrm{e} + \varepsilon^{-1}))^{-1}$.

Now define the auxiliary process $(\overline{\vartheta}_t^\varepsilon)_{t\geq 0}$ by $\overline{\vartheta}_t^\varepsilon = \vartheta - \vartheta_t^\varepsilon$ and introduce $(A_t^\varepsilon)_{t\geq 0}$ and $(B_t^\varepsilon)_{t\geq 0}$ by

$$A_t^\varepsilon = a + \overline{\vartheta}_t^\varepsilon \cdot \chi_{\{|\overline{\vartheta}_t^\varepsilon| \leq \gamma_t^\varepsilon\}}, \qquad B_t^\varepsilon = \overline{\vartheta}_t^\varepsilon \cdot \chi_{\{|\overline{\vartheta}_t^\varepsilon| > \gamma_t^\varepsilon\}}, \quad t \geq 0,$$

where $\chi_{\{a<b\}} = 1$ if $a < b$ and 0 otherwise.

By the definition, for the processes (Δ_t^ε) and (δ_t^ε) the following equations hold

$$d\Delta_t^\varepsilon = A_t^\varepsilon \Delta_t^\varepsilon dt + b_t^\varepsilon x_t^0 dt, \quad t \geq 0, \tag{17}$$

where $b_t^\varepsilon = \overline{\vartheta}_t^\varepsilon \cdot \chi_{\{|\overline{\vartheta}_t^\varepsilon| \leq \gamma_t^\varepsilon\}}$, and

$$d\delta_t^\varepsilon = A_t^\varepsilon \delta_t^\varepsilon dt + B_t^\varepsilon x_t^\varepsilon dt, \quad t \geq 0. \tag{18}$$

Note, that according to the condition (16) and the definitions of (A_t^ε) and $(\vartheta_t^\varepsilon)$, the process (A_t^ε) is, for $\varepsilon \leq \varepsilon_0^*$ uniformly bounded from above

$$A_t^\varepsilon \leq -a^\varepsilon, \quad t \geq 0,$$

where $a^\varepsilon = -(a + \pi(\varepsilon))$ is positive. As follows, the fundamental function $\Psi_\varepsilon(s,t) = \mathrm{e}^{\int_s^t A_l^\varepsilon dl}$ of the processes (17) and (18) can be estimated from above:

$$\Psi_\varepsilon(s,t) \leq \mathrm{e}^{-a^\varepsilon(t-s)}, \quad 0 \leq s \leq t. \tag{19}$$

To prove (7) and (8) we can derive, using (19), estimators of the second moments of the processes (Δ_t^ε) and (δ_t^ε).

Define for every $\theta \geq 1$ the numbers

$$\overline{C}_x^0(\theta) = \sup_{t\geq 0} E(\overline{x}_t^0)^{2\theta}, \quad C_x^0(\theta) = \sup_{t\geq 0} E(x_t^0)^{2\theta},$$

$$C(\theta,\varepsilon) = \frac{2}{\mu_\varepsilon^4 a^\varepsilon} \cdot (\frac{2}{3} C_1^{\frac{2}{\theta}}(\theta) + C_2^{\frac{2}{\theta}}(\theta)),$$

$$C_1(\theta) = \frac{1}{2}\left(\frac{32(\vartheta_1 - \vartheta_0)\theta}{a^2}\right)^\theta \cdot \overline{C}_x^0(\theta),$$

$$C_2(\theta) = \frac{1}{4}\left(\frac{4(\vartheta_1 - \vartheta_0)}{a^2}\right)^\theta \cdot (\overline{C}_x^0(2\theta) + E x_0^{4\theta}),$$

$$f_0 = -\frac{1}{2(a + \vartheta - \vartheta_1)}, \quad \mu_\varepsilon = f_0 \cdot \pi(\varepsilon),$$

$$\tilde{t}_\varepsilon = [c_{0,\varepsilon} \cdot (f_0 - \mu_\varepsilon)^{-1}] \vee [8(a^\varepsilon)^{-1}],$$

where $c_{0,\varepsilon}$ is a special chosen number;

for $t \geq \tilde{t}_\varepsilon$ and $\theta > 2$ the function

$$\Delta(\tilde{t}_\varepsilon, t, \theta) = \frac{1}{a^\varepsilon} \cdot (C_x^0(\frac{\theta}{\theta - 2}))^{\frac{\theta-2}{\theta}} \cdot$$

$$\cdot\{\frac{1}{a^\varepsilon}(\vartheta_1 - \vartheta_0)^2 \mathrm{e}^{-a^\varepsilon(t - \tilde{t}_\varepsilon)} + \frac{C(\theta,\varepsilon)}{t^2}\}.$$

Set the number

$$\lambda_0 = 8\mathrm{e} f_0^{-1} a^{-2} \cdot (E x_0^2 - \frac{1}{2a}) \tag{20}$$

and define the time

$$t_\varepsilon = \inf\{t \geq \tilde{t}_\varepsilon : \ \Delta(\tilde{t}_\varepsilon, t, 2\pi^{-1}(\varepsilon)) \leq \pi^2(\varepsilon) \cdot \varepsilon\}. \tag{21}$$

Theorem 2.1: Let u^ε be a control law of the object (3) as well as the number λ_0 and the time t_ε are defined by the formulae (13), (20) and (21) respectively.

Then for the object (3) for ε small enough the relations

$$\sup_{t\geq 0} \ E(x_t^\varepsilon)^2 \leq L$$

and

$$\sup_{t\geq t_\varepsilon} \ E(x_t^\varepsilon - x_t^0)^2 \leq \varepsilon$$

are fulfilled. The time t_ε has the following rate of increase

$$\varepsilon \cdot [t_\varepsilon - \lambda_0 \cdot \varepsilon^{-1} \ln \varepsilon^{-1}] = O(1) \quad as \ \varepsilon \to 0. \tag{22}$$

Corollary 2.1: Using the definition of the stopping time $\tau_{0,\varepsilon}$ and Theorem 2.1 the following limiting relation can be proved:

$$\lim_{\varepsilon\to 0} \ \frac{\tau_{0,\varepsilon}}{t_\varepsilon} = 1 \quad a.s.$$

Example 2.1: The sequences $(\gamma_{n,\varepsilon})$ and $(c_{n,\varepsilon})$ can be taken as follows:

$$\gamma_{n,\varepsilon} = \gamma_{0,\varepsilon} \cdot \chi_{\{n=0\}} + \gamma_{1,\varepsilon} \cdot n^{-\rho} \cdot \chi_{\{n\geq 1\}},$$

$$c_{n,\varepsilon} = c_{0,\varepsilon} \cdot \chi_{\{n=0\}} + c_{1,\varepsilon} \cdot n^{\alpha} \cdot \chi_{\{n\geq 1\}},$$

where $\gamma_{0,\varepsilon}$, $\gamma_{1,\varepsilon}$, $c_{0,\varepsilon}$ and $c_{1,\varepsilon}$ are some special chosen numbers and $\alpha > 2\rho > 0$.

Remark 2.1: The time $\tilde{t}_\varepsilon$ depends on the unknown parameter ϑ and can be estimated from above, using the inequality $f_0 \geq \dfrac{1}{2(\vartheta_1 - \vartheta_0 - a)}$, by the known time $\bar{t}_\varepsilon \geq \tilde{t}_\varepsilon$ defined as

$$\bar{t}_\varepsilon = [2(\vartheta_1 - \vartheta_0 - a)c_{0,\varepsilon} \cdot (1 - \pi(\varepsilon))^{-1}] \vee [8(a^\varepsilon)^{-1}].$$

Thus the relation (8) for the object (3) is fulfilled for the known time t_ε defined in (21) with $\bar{t}_\varepsilon$ instead of $\tilde{t}_\varepsilon$.

In this case the following limiting equality holds true:

$$\varepsilon \cdot [t_\varepsilon - \lambda_1 \varepsilon^{-1} \ln \varepsilon^{-1}] = O(1) \quad as\ \varepsilon \to 0,$$

where λ_1 is a known number:

$$\lambda_1 = 16(\vartheta_1 - \vartheta_0 - a)ea^{-2} \cdot (Ex_0^2 - \frac{1}{2a}).$$

B. Construction of the control law for Problem II

Consider the process of the type (9)

$$dx_t^\varepsilon = \vartheta x_{t-r}^\varepsilon dt + u_t^\varepsilon dt + dw_t, \quad t \geq 0$$

with a control function u^ε, for which we shall prove that (7) and (8) hold, defined as follows.

It should be noted that solutions of SDDE's have essentially more complicate asymptotic behavior in comparison with linear SDE's, see [9]. Thus, in particular, the adaptive control algorithm, constructed similar to the previous section II-A for the system without delay is more difficult for investigation.

Let $(\tilde{c}_{n,\varepsilon})_{n\geq 0}$ be an unboundedly increasing sequence of positive numbers, satisfying some general conditions and $\tilde{\tau}_\varepsilon = (\tilde{\tau}_{n,\varepsilon})_{n\geq -1}$ be the sequence of stopping times

$$\tilde{\tau}_{n,\varepsilon} = \inf\{T > 0 : \int_0^T (x_{t-r}^\varepsilon)^2 dt = \tilde{c}_{n,\varepsilon}\},\ n \geq 0,\ \tilde{\tau}_{-1,\varepsilon} = 0.$$

Define the sequence $(\tilde{\vartheta}_{n,\varepsilon})_{n\geq 0}$ of sequential estimators of ϑ by

$$\tilde{\vartheta}_{n,\varepsilon} = \frac{1}{\tilde{c}_{n,\varepsilon}} \int_0^{\tilde{\tau}_{n,\varepsilon}} x_{t-r}^\varepsilon (dx_t^\varepsilon - u_t^\varepsilon dt), \quad n \geq 0$$

and the sequence $(\vartheta_{n,\varepsilon})_{n\geq 0}$:

$$\vartheta_{n,\varepsilon} = (\tilde{\vartheta}_{n,\varepsilon} \wedge \vartheta_1) \vee \vartheta_0, \quad n \geq 0.$$

Define $\tilde{b}(\vartheta) = b + \vartheta - \overline{\vartheta}$, $\overline{\vartheta} = \dfrac{\vartheta_0 + \vartheta_1}{2}$. The real number $\tilde{a}^\varepsilon$ will be chosen in such a way, that the equation

$$dx_t^\varepsilon = \tilde{a}^\varepsilon x_t^\varepsilon dt + \tilde{b}(\vartheta) x_{t-r}^\varepsilon dt + dw_t, \quad 0 \leq t < \tilde{\tau}_{0,\varepsilon} \qquad (23)$$

admits a stationary solution for all $\vartheta \in [\vartheta_0, \vartheta_1]$ (and that moreover we have $(\tilde{a}^\varepsilon, \tilde{b}(\vartheta)) \to (a, b)$ if $\varepsilon \to 0$ and $\vartheta_1 - \vartheta_0 \to 0$). The condition on $\tilde{a}^\varepsilon$ is difficult because the set of parameters (a, b), where the equation

$$dx_t = ax_t dt + bx_{t-r} dt + dw_t \qquad (24)$$

has a stationary solution is more complicate than in the Ornstein-Uhlenbeck case I (see, e.g. [9]).

In [9] the region of all parameters (a, b), such that the process (24) with $r = 1$ is stable, is characterized by using a function $u(a)$, $a < 1$, defined as follows:
introduce a parametric curve $(a(\xi), b(\xi)), \xi > 0, \xi \neq \pi, 2\pi, \ldots$ in R^2 by

$$a(\xi) = \xi \cot \xi,\ b(\xi) = -\xi / \sin \xi,$$

then $b = u(a)$ is the branch of this curve corresponding to $\xi \in (0, \pi)$.

Define the numbers $\beta_1 = b - \dfrac{\vartheta_1 - \vartheta_0}{2}$, $\beta_2 = b + \dfrac{\vartheta_1 - \vartheta_0}{2}$ and the number α for the case $\beta_1 \leq r^{-1}$ as a solution of the equation $u(\alpha r) = \beta_1 r$.

Now we are ready to define the parameter $\tilde{a}^\varepsilon$ as follows:

$$\tilde{a}^\varepsilon = \begin{cases} a \cdot \chi_{\{\alpha - \pi(\varepsilon) \geq a\}} + (\alpha - \pi(\varepsilon)) \cdot \chi_{\{\alpha - \pi(\varepsilon) < a\}}, \\ \beta_2 \leq r^{-1}, \\ a \cdot \chi_{\{\alpha \wedge \beta_2 - \pi(\varepsilon) \geq a\}} + (\alpha \wedge \beta_2 - \pi(\varepsilon)) \cdot \\ \chi_{\{\alpha \wedge \beta_2 - \pi(\varepsilon) < a\}}, \quad \beta_1 < r^{-1} < \beta_2, \\ a \cdot \chi_{\{\beta_2 - \pi(\varepsilon) \geq a\}} + (\beta_2 - \pi(\varepsilon)) \cdot \chi_{\{\beta_2 - \pi(\varepsilon) < a\}}, \\ \beta_1 \geq r^{-1}. \end{cases}$$

It is clear, that (23) with such parameters has a stationary solution.

For the defined function $\tilde{a}^\varepsilon$ and the number $\overline{\vartheta}$, we introduce the piecewise-constant functions $(a_t^\varepsilon)_{t\geq 0}$ and $(\vartheta_t^\varepsilon)_{t\geq -r}$ by

$$a_t^\varepsilon = \begin{cases} \tilde{a}^\varepsilon, & 0 \leq t < \tilde{\tau}_{0,\varepsilon}, \\ a, & t \geq \tilde{\tau}_{0,\varepsilon}; \end{cases}$$

$$\vartheta_{t-r}^\varepsilon = \begin{cases} \overline{\vartheta}, & 0 \leq t < \tilde{\tau}_{0,\varepsilon}, \\ \vartheta_{n,\varepsilon}, & \tilde{\tau}_{n,\varepsilon} \leq t < \tilde{\tau}_{n+1,\varepsilon}, \quad n \geq 0. \end{cases}$$

As the control law u^ε we now define

$$u_t^\varepsilon = a_t^\varepsilon x_t^\varepsilon + (b - \vartheta_{t-r}^\varepsilon) x_{t-r}^\varepsilon, \quad t \geq 0. \qquad (25)$$

Similar to the previous section the main idea is to construct an appropriate decomposition (26) below for the object (9) with the control process (25).

Define auxiliary processes $(A_t^\varepsilon)_{t\geq 0}$, $(b_t^\varepsilon)_{t\geq 0}$ and $(B_t^\varepsilon)_{t\geq 0}$ as follows

$$A_t^\varepsilon = b + b_t^\varepsilon,\ b_t^\varepsilon = \overline{\vartheta}_{t-r}^\varepsilon \chi_{\{|\overline{\vartheta}_{t-r}^\varepsilon| \leq \tilde{\gamma}_t^\varepsilon\}},\ B_t^\varepsilon = \overline{\vartheta}_{t-r}^\varepsilon \chi_{\{|\overline{\vartheta}_{t-r}^\varepsilon| > \tilde{\gamma}_t^\varepsilon\}},$$

where $\overline{\vartheta}_t^\varepsilon = \vartheta - \vartheta_t^\varepsilon$, $t \geq -r$ and $(\tilde{\gamma}_t^\varepsilon)_{t\geq 0}$ is a piecewise constant function defined similar to (15).

Now we write the equation (9) with the control (25) in the form

$$dx_t^\varepsilon = a_t^\varepsilon x_t^\varepsilon dt + A_t^\varepsilon x_{t-r}^\varepsilon dt + B_t^\varepsilon x_{t-r}^\varepsilon dt + dw_t, \quad t \geq 0.$$

Define the auxiliary process $(\overline{a}_t^\varepsilon)_{t\geq 0}$, $\overline{a}_t^\varepsilon = a_t^\varepsilon - a$ as well as processes $(y_t^\varepsilon)_{t\geq -r}$, $(\Delta_t^\varepsilon)_{t\geq -r}$ and $(\delta_t^\varepsilon)_{t\geq -r}$, satisfying the equations:

$$dy_t^\varepsilon = a_t^\varepsilon y_t^\varepsilon dt + A_t^\varepsilon y_{t-r}^\varepsilon dt + dw_t, \quad y_s^\varepsilon = x_0(s), \quad s \in [-r, 0],$$

$$\Delta_t^\varepsilon = y_t^\varepsilon - x_t^0, \quad \Delta_s^\varepsilon = 0, \quad s \in [-r, 0],$$

$$\delta_t^\varepsilon = x_t^\varepsilon - y_t^\varepsilon, \quad \delta_s^\varepsilon = 0, \qquad s \in [-r, 0], \quad t \geq 0.$$

Then the observed process x^ε, similar to the Problem I can be represented in the form

$$x_t^\varepsilon = x_t^0 + \Delta_t^\varepsilon + \delta_t^\varepsilon, \quad t \geq -r. \qquad (26)$$

By the definition, the processes (Δ_t^ε) and (δ_t^ε) satisfy the following equations

$$d\Delta_t^\varepsilon = a_t^\varepsilon \Delta_t^\varepsilon dt + A_t^\varepsilon \Delta_{t-r}^\varepsilon dt + \overline{a}_t^\varepsilon x_t^0 dt + b_t^\varepsilon x_{t-r}^0 dt,$$

$$d\delta_t^\varepsilon = a_t^\varepsilon \delta_t^\varepsilon dt + A_t^\varepsilon \delta_{t-r}^\varepsilon dt + B_t^\varepsilon x_{t-r}^\varepsilon dt, \quad t \geq 0$$

and are stable.

To prove (7) and (8) we can derive, similar to the previous section, estimators of the second moments of the processes (Δ_t^ε) and (δ_t^ε).

Similar to Problem I, we have defined a function t_ε^* and a number λ_0^* satisfying the following relation:

$$\varepsilon \cdot [t_\varepsilon^* - \lambda_0^* \cdot \varepsilon^{-1} \ln \varepsilon^{-1}] = O(1) \quad \text{as} \quad \varepsilon \to 0. \tag{27}$$

Theorem 2.2: For the constructed u^ε for some number L^* independent from ε it holds

$$\sup_{t \geq 0} E(x_t^\varepsilon)^2 \leq L^*$$

and for ε small enough

$$\sup_{t \geq t_\varepsilon^*} E(x_t^\varepsilon - x_t^0)^2 \leq \varepsilon.$$

III. Summary

This paper presents a certainty equivalence design method with application for two continuous-time stochastic systems with unknown parameters. The main aim is to approximate the observed processes by choosing the control process $(u_t)_{t \geq 0}$ to the stable Ornstein-Uhlenbeck process (Problem I) and to the stable related process satisfying SDDE (Problem II) in the sense of the inequality (8). In comparison with the often used as a criterion limit target inequality (2), we have found the time t_ε (ε is a threshold quantity) after that our target inequality (8) is fulfilled. It is shown that the time t_ε in (8) has equal rates of increase $\varepsilon^{-1} \ln \varepsilon^{-1}$ in both problems. Moreover, the target inequality (7), which ensures the stability of the controlled processes is established.

In Problem I the time t_ε agrees closely with the time of obtaining of the first estimator with a given accuracy of the unknown parameter ϑ of the object for ε small enough. Moreover, in Problem I this time has known upper bound and in Problem II the limiting constant λ_0^* (see (27)) can be estimated.

The constructed control law works in a real time. Simultaneously the problem of estimation with guaranteed in the sense (12) accuracy of the unknown parameter ϑ in both the problems is solved.

The method presented in the paper can be easily applied for the special type multidimensional systems considered, e.g. in [25]. Similar problem for more general controlled multidimensional systems can be solved using sequential estimators of unknown parameters, presented in [10].

Acknowledgement

We are thankful to RFBR – DFG 10-01-91330 Grant for support of our research.

References

[1] K.J. Åström, G.C. Goodwin and P.R. Kumar (Editors), *Adaptive Control, Filtering, and Signal Processing*, Springer-Verlag, New York Inc., 1995.

[2] Jinde Cao, Jun Wang, *Delay-dependent robust stability of uncertain nonlinear systems with time delay*, Applied Mathematics and Computation, 154(1), 25 June (2004), pp. 289–297.

[3] Carlen, E., Kree, P., *L^p estimates on iterated stochastic integrals*, Ann. Probab. 19 (1991); pp. 354–368.

[4] Patrizio Colaneri, Jos C. Geromel, Arturo Locatelli, Jos C. Geromel and Arturo Locatelli, *Control Theory and Design*, Elsevier Ltd., 1997.

[5] T.E. Duncan and B. Pasik-Duncan, *Adaptive control of linear delay time systems*, Stochastics, 24 (1988), pp. 45–74.

[6] T.E. Duncan, B. Pasik-Duncan and L. Stettner, *On the ergodic and the adaptive control of stochastic differential delay systems*, J. Optimization Theory and Application. 81(3) (1994), pp. 509–531.

[7] Fomin V.N., Fradkov A.L., Jakubovich V.A., *Adaptive control of Dynamic Objects*, Moscow: Nauka, 1981.

[8] Fradkov A.L., Miroshnik I.V., Nikiforov V.O., *Nonlinear and adaptive control of complex systems*, Dordrecht: Kluwer Academic Publishers, 1999.

[9] A. A. Gushchin, and U. Küchler, *Asymptotic inference for a linear stochastic differential equation with time delay*, Bernoulli, 5(6) (1999), pp. 1059–1098.

[10] U. Küchler and V.Vasiliev, *On guaranteed parameter estimation of a multiparameter linear regression process*, Automatica, 46(4), April (2010), pp. 637–646.

[11] J. S. Luo, A. Johnson, P. P. J. van den Bosch, *Delay-independent robust stability of uncertain linear systems*, Systems and Control Letters, 24(1), 9 January (1995), pp. 33–39.

[12] Driss Mehdi, Mohammed Al Hamid, Franois Perrin, *Robustness and optimality of linear quadratic controller for uncertain systems*, Automatica, 32(7), July (1996), pp. 1081–1083.

[13] D.E. Miller, *Adaptive stabilization using a nonlinear time-varying controller*, IEEE Trans. Automatic Control, 39(7) (1994), pp. 1349–1359.

[14] Emmanuel Moulay, Wilfrid Perruquetti, *Finite time stability and stabilization of a class of continuous systems*, Journal of Mathematical Analysis and Applications, 323(2), 15 November (2006), pp. 1430–1443.

[15] Gevorg Nahapetian, Wei Ren, *Uncertainty structures in adaptive and robust stabilization*, Automatica, 31(11) (1995), pp. 1565-1575.

[16] Nemirovski A.S., Tsypkin Ya. Z., *On optimal algorithms of automatic control*, Automat. and Remote Control, 12 (1984), pp. 64–77.

[17] Ian R. Petersen, B. Ross Barmish, *Control effort considerations in the stabilization of uncertain dynamical systems*, Systems and Control Letters, 9(5), November (1987), pp. 417–422.

[18] K.J. Roger, W. Brockett, *Reduced Complexity Control Systems*, Proceedings of the 17th World Congress The International Federation of Automatic Control Seoul, Korea, July 6-11 (2008), pp. 1–6.

[19] V.G. Sragovich, *Mathematical theory of adaptive control. Interdisciplinary Mathematikal Sciences*, Vol. 4, World Scientific Publishing Co. Pte. Ltd., 2006.

[20] Taro Tsujino, Takao Fujii, Kehui Wei, *On the connection between controllability and stabilizability of linear systems with structural uncertain parameters*, Automatica, 29(1), January (1993), pp. 7–12.

[21] V.A.Vasiliev, *Nonparametric Estimation of Errors Distribution in Linear Stochastic Control Systems. - 'Magdeburger Stochastic Tage-2002'*, Materials of German Open Conference on Probability and Statistics, Proc. of Conf., Germany, Magdeburg, 19-22 March, 2002.

[22] V.A.Vasiliev, A.V.Dobrovidov, G.M.Koshkin, *Nonparametric estimation of functionals of stationary sequences distributions*, Moscow.: Nauka, 2004 (in Russian).

[23] V.A.Vasiliev, V.V.Konev, *On Optimal Adaptive Control of a Linear Stochastic Process*, Proc.15th IMACS World Congress on Scientific Computation, Modelling and Applied Mathematics, Berlin, Germany, August 24-29, 5, (1997), pp. 87–91.

[24] Vasiliev V.A., Koshkin G.M., *Nonparametric identification of autoregressions*, Theory Probab. Appl., 43(3) (1999), pp. 507–517.

[25] Vorobeichikov S.E., Konev V.V., *On sequential identification of stochastic systems*, Tehnical Cybernetics, 4 (1980), pp. 64–77.

[26] Yao-Feng Ren, *On the Burkholder-Davis-Gundy inequalities for continuous martingales*, Statistics and Probability Letters, 78 (2008), pp. 3034–3039.

19 Fractional Order Ultra Low-Speed Position Servo: Improved Performance via Describing Function Analysis

Fractional Order Ultra Low-Speed Position Servo: Improved Performance via Describing Function Analysis

Ying Luo*†, YangQuan Chen† and YouGuo Pi‡
*Dept. of Automation Science and Engineering, South China University of Technology, Guangzhou, P. R. China
Email: ying.luo@ieee.org
† Center for Self-Organizing and Intelligent Systems (CSOIS),
Dept. of Electrical and Computer Engineering, Utah State University, 4120 Old Main Hill, Logan, UT 84322, USA.
Email: yqchen@ieee.org
‡Dept. of Automation Science and Technology, South China University of Technology, Guangzhou, P. R. China
Email: auygpi@scut.edu.cn

***Abstract*—In [1][2], a new systematic design method for fractional order proportional and derivative (FOPD) controller is proposed for a class of typical second-order plants. Simulation and experimental results show that the dynamic performance and robustness of the position ramp response at normal speed with the designed FOPD controller outperforms that with the ITAE optimized traditional integer order proportional and integral (IOPI) controller. Furthermore, we found that, for the ultra low speed position tracking with significant friction effect in the same experimental system in [2], the tracking performance using the designed FOPD controller is much better than that using the optimized IOPI controller. In this paper, using the describing function method and the Bode plots analysis, the observed advantage of the designed FOPD controller for the nonlinear low speed position tracking system with friction effect over the optimized IOPI controller is explained, which is consistently demonstrated by our extended experimental results.**

I. Introduction

Friction is the force resisting the relative lateral motion of solid surfaces, fluid layers, or material elements in contact. This common nonlinear phenomenon has an universal impact in all regimes of operation in mechanisms and produces undesirable behaviors in control systems such as the tracking errors, and limit cycles [3]. Especially, in the high-precision position control systems, the performance is inherently affected by the friction effect. Compensation for the friction and attenuation of its effects has been addressed in many papers over the years [3], [4], [5], [6]. The describing function method (DF) is widely used as a common tool for the nonlinear system analysis [7], [8], [9], [10], [11]. In [6], using the describing function method, different approaches for the prediction of limit cycles in control systems with friction are discussed based on a simple stick-slip motion example. The existing limit cycles cannot be predicted using the describing function for only the friction nonlinearity part. But the describing function combining part of the plant with the friction model can capture the behavior of the friction with zero velocity, and the limit cycles can be predicted. However, this describing function depends on three parameters, the amplitude, frequency and offset, comparing with the normal two parameters, the amplitude and frequency.

On the other hand, the application of fractional calculus attracts increasing attentions in control domain in recent years [12], [13], [14]. It is remarkable to see the increasing number of studies related to the theory and applications of fractional order controllers (FOC), especially, the fractional order PID controller. In [1][2], a new systematic design method for fractional order proportional and derivative (FOPD) controller is proposed for a class of typical second-order plants. The tuned FOPD controller can ensure that the given gain crossover frequency and phase margin are fulfilled, and the phase derivative w. r. t. the frequency is zero, i.e., phase Bode plot is flat, around the given gain crossover frequency. So that the closed loop system is robust to gain variations. Simulation and experimental results show that the dynamic performance and robustness of the position ramp response at normal speed with the designed FOPD controller outperforms that with the ITAE optimized traditional integer order proportional and integral (IOPI) controller. Furthermore, we found that, for the low speed position tracking with significant friction effect in the same experimental platform in [1][2], the tracking performance using the designed FOPD controller is much better than that using the optimized IOPI controller. Based on this favorable experimental phenomenon, the theoretical analysis is needed for the clear understanding.

In this paper, using the describing function method and the Bode plots analysis, the observed advantage of the designed FOPD controller for the nonlinear low speed position tracking system with friction effect over the optimized IOPI controller is explained, which is consistently demonstrated by our extended experimental results.

II. Low Speed Position Tracking using FOPD and IOPI

In this section, the main idea of the FOPD controller design and the performance comparison with the ITAE optimized IOPI controller in [1][2] are introduced briefly. The experiment comparison for the low speed position tracking with significant friction effect using the designed FOPD and the optimized IOPI are presented.

A. Introduction to the FOPD Design of the Position Tracking without Considering the Friction Effect

A new systematic design method for FOPD controller is proposed for a class of typical second-order plants without

considering the friction effect in [1][2]. The key points of this FOPD controller systematic design scheme are that the designed FOPD controller can ensure that the given gain crossover frequency and phase margin are fulfilled, and furthermore the phase derivative w. r. t. the frequency is zero. So that, the closed loop system is robust to gain variations.

For FOPD controller design, a class of second-order plants $P(s)$ is described by (1),

$$P(s) = \frac{K}{s(Ts+1)}, \tag{1}$$

which can approximately model a DC motor position servo system. The experimental platform of the dynamometer for the position tracking in [2] was identified as the second-order system as (1) with $K = 1.52$ and $T = 0.4$.

The FOPD controller has the following form of transfer function,

$$C(s) = K_p(1 + K_d s^{\mu}), \tag{2}$$

where $\mu \in (0, 1]$.

Three specifications are proposed to design the FOPD controller.

(i) Phase margin specification

$$\mathrm{Arg}[G(j\omega_c)] \quad = \quad \mathrm{Arg}[C(j\omega_c)P(j\omega_c)] = -\pi + \phi_m,$$

where the ϕ_m and ω_c are the desired phase margin and gain crossover frequency, respectively.

(ii) Robustness specification to the plant gain variations

$$(\frac{\mathrm{d}(\mathrm{Arg}(C(j\omega)P(j\omega)))}{\mathrm{d}\omega})_{\omega=\omega_c} = 0,$$

with the condition that the phase Bode plot is flat, around the gain crossover frequency. It means that the system is more robust to gain changes and the overshoots of the response are almost the same.

(iii) Gain crossover frequency specification

$$|G(j\omega_c)|_{dB} = |C(j\omega_c)P(j\omega_c)|_{dB} = 0.$$

With these specifications, the gain crossover frequency is set as $\omega_c = 10(rad/s)$, and the desired phase margin is set as $\Phi_m = 70^{\circ}$. Moreover, the robustness to gain variations is required. Using the dynamometer experimental model, according to the numerical method in [1], we can obtain the parameters of the FOPD controller as $\mu = 0.844$, $K_d = 0.368$ and $K_p = 13.860$. Meanwhile, the parameters of the ITAE optimized IOPI controller are designed as $K_p = 2.6531$ and $K_i = 1.1662$ [1][15].

Simulation and experimental results show that, the dynamic performance and robustness of the position ramp response at normal speed with the designed FOPD controller outperforms that with the ITAE optimized IOPI controller [1][2].

B. Low Speed Position Tracking Performances with the FOPD and IOPI Controllers

In the same experimental plant, and using the same designed FOPD and optimized IOPI controllers as in [1], if the position ramp is generated by the integration of a normal reference speed without considering the real friction except the viscous part, the experimental system can be described in Fig. 1, which is equal to the system in Fig. 2. If both the nonlinear unmodelled friction effect and the viscous friction are considered in the system model, the closed-loop system can be shown in Fig. 3 with C as the designed FOPD or the optimized IOPI.

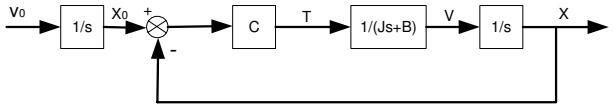


Fig. 1. Position tracking control diagram with constant speed reference without considering friction

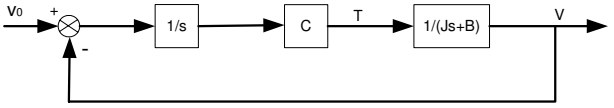


Fig. 2. Position tracking control equivalent diagram with constant speed reference without considering friction

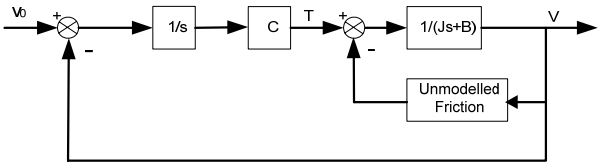


Fig. 3. Position tracking control equivalent diagram with constant speed reference and friction

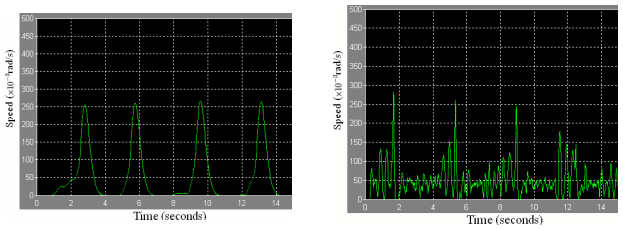
When the reference speed used for the position ramp tracking is reduced to a very small value as $0.05\ rad/s$, then, the friction effect is significant and not negligible. So, Fig. 3 should be used to describe the closed-loop experimental system. Figs. 4(a), 4(b), 5(a) and 5(b) show speed and position output of the position tracking with constant ultra low-speed reference, it is obvious that, the tracking performance using the designed FOPD in Fig. 4(b) is much better than that using the optimized IOPI in Fig. 4(a).

III. Static / Dynamic Models of Friction and Describing Functions for Friction Model

In this section, the different friction models and two uncoupling methods of the linear part and nonlinear part are presented.

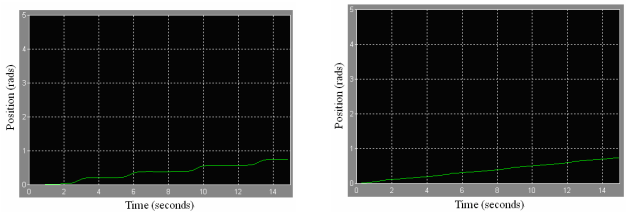
A. Static and Dynamic Models of Friction

In general, the friction models are described by a discontinuous relation between the relative velocity in between the surfaces contacted and the resulting friction force. The friction force can be briefly divided into the traditional static models which are expressed by the static equations and the combinations of coulomb friction, viscous friction and so on [16]. The friction force can be also be modeled by the dynamic models proposed in the last few decades with the



(a) IOPI (b) FOPD

Fig. 4. Speed output of position tracking with constant speed reference using optimized IOPI and designed FOPD



(a) IOPI (b) FOPD

Fig. 5. Position output of position tracking with constant speed reference using optimized IOPI and designed FOPD

differential equations [5][17]. Many models are defined without considering the velocity zero. For this particular velocity, the friction force depends on the applied force. As presented in [5], the proposed LuGre dynamic friction model combines the stiction behavior, i.e., the Dahl effect, with arbitrary steady state friction characteristics which can include the Stribeck effect and the zero velocity friction. The typical LuGre model is useful for various control tasks, and is given by [5],

$$\dot{z} = v - \frac{|v|}{g(v)}z, \tag{3}$$

$$F = \sigma_0 z + \sigma_1 \dot{z} + \sigma_2 v, \tag{4}$$

$$\sigma_0 g(v) = F_C + (F_S - F_C)e^{-(v/v_s)^2}, \tag{5}$$

where the average deflection of the bristles is denoted by z; v is the relative velocity between the two surfaces; the function g is positive and depends on many factors such as material properties, lubrication and temperature; σ_0 is the stiffness, and σ_1, σ_2 are damping coefficients; F_C is the Coulomb friction level; F_S is the level of the stiction force, and v_s is the Stribeck velocity.

B. Describing Functions for Friction Models and Two Uncoupling Methods of Linear and Nonlinear Parts

The experimental platform for the position tracking with consideration of the friction effect is shown in Fig. 3. Using the describing function method, the transfer function of the nonlinear block is described by the relationship between the output response $y(t)$ and the frequency ω. In the output response, only the first harmonic $y_1(t)$ with the same frequency as that in the input signal is considered,

$$y_1(t) = a\cos\omega t + b\sin(\omega t) = c\sin(\omega t + \varphi), \tag{6}$$

where

$$\begin{aligned} a &= \frac{2}{\pi}\int_0^\pi y(t)\cos(\omega t)\mathrm{d}(\omega t), \\ b &= \frac{2}{\pi}\int_0^\pi y(t)\sin(\omega t)\mathrm{d}(\omega t), \\ c &= \sqrt{a^2+b^2}, \quad \varphi = \arctan(b/a). \end{aligned} \tag{7}$$

The describing function of the nonlinear block can be expressed by the gain and phase shift between the first harmonic of the output and the sinusoid input as below,

$$N_I(A,\omega) = \frac{c}{A}e^{j\varphi}. \tag{8}$$

In order to use the describing function method to analysis the system performance, we need to get the approximation closed-loop system with uncoupled linear and nonlinear parts as in Fig. 6, from the position tracking system with friction nonlinear effect in Fig. 3. There are two methods for uncoupling the linear and nonlinear parts as shown in Fig. 7(a) and Fig. 7(b). The straightforward way is to treat only the friction as the nonlinear part, and the other items as the linear part in Fig. 7(a). Following the analysis in [6], with the simple example of the stick-slip motion, it can be seen that, the velocity is the input and the friction is the output for the describing function analysis.Then, the nonlinear part will not be affected by the behavior of the friction force with exact zero velocity. So, the intricate behavior of the friction with zero velocity is neglected unfortunately. Meanwhile, the Nyquist curves for the linear and nonlinear parts are plotted in [6]. It is figured out that there is no intersection which the limit cycle frequency can be obtained from, for the two curves except the origin point. So, the analysis with the uncoupling method in Fig. 7(a) doesn't predict any limit cycle [6] in spite of the average velocity v_0, frequency ω and amplitude A.

For the other uncoupling method in Fig. 7(b), the force T is the input and the velocity is the output of the nonlinear part which includes not only the friction but also the system dynamic G. In this case, the friction force has the possibility to counteract the applied force and keep the velocity as zero. During one period of the sinusoid input, sticking may occur. Therefore, the essential characteristics of the friction can be captured in the describing function of the nonlinear part in Fig. 7(b). This time, a mean value of the force T have to be included as one of the input parameters for the nonlinear block [6]. So, this method in Fig. 7(b) are chosen to uncouple the linear and nonlinear parts of our experimental nonlinear system. The describing function will hence depend on three parameters, the amplitude A, the frequency ω and the mean force T_0. The output will be an oscillation with a mean value different from zero. The describing function under this case should be notated as below,

$$N(A,\omega,T_0) = [|N|e^{\phi}, v_a],$$

where the $|N|$ and ϕ are the gain and phase shift of the describing function, v_a is the average value of the output velocity of the nonlinear part.

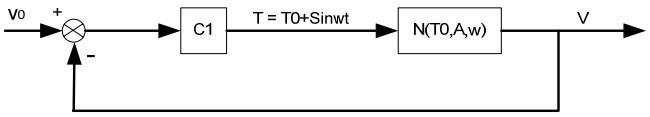
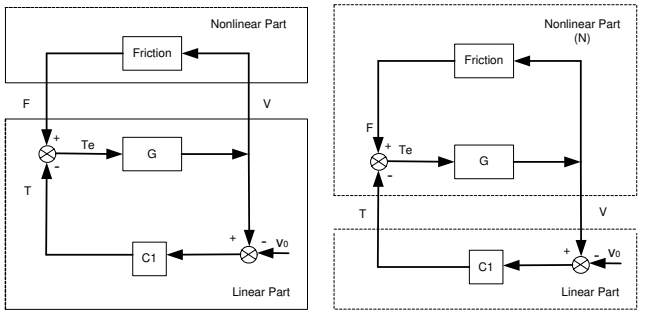


Fig. 6. Approximation closed-loop system with linear and nonlinear parts



(a) Nonlinear part only with friction (b) Nonlinear part with friction and system dynamic

Fig. 7. Two methods of uncoupling linear and nonlinear parts

From Figs. 4 and 5, we can see the output signals of the nonlinear part, using the uncoupling method in Fig. 7(b) with IOPI or FOPD controller, for the low speed position tracking in the example introduced in Sec. II. It can be seen that, both the outputs with IOPI and FOPD have limit cycles.

IV. Bode Plot Comparison with IOPI and FOPD Controllers Using Describing Function

In order to reveal the potential advantage of the designed FOPD controller for the nonlinear low speed position tracking system with friction effect over the optimized IOPI controller, the Bode plot analysis with the describing function is used.

First, the Bode plots of open-loop transfer functions are drawn in Fig. 8 using IOPI and FOPD without considering friction, for the position tracking system. It can be seen that, when the frequency around the gain crossover frequency, and ω>$0.2rad/sec$, the amplitude with FOPD is bigger than that with IOPI. Meanwhile, the phase delay with FOPD is much smaller than that with IOPI, which means that the relative stability of the system is significant improved by the designed FOPD controller comparing with the optimized IOPI controller.

Second, for the block diagram as shown in Fig. 7(b), the mean forces T_{0IOPI} and T_{0FOPD} of the input signals T of the nonlinear block with IOPI and FOPD are measured as 0.577 and 0.622, respectively. So, the describing function of the nonlinear part can be calculated following the method introduced in Sec. III-B. Using IOPI with $T_{0IOPI} = 0.577\ N*m$, the 3D/2D Bode plots of the amplitudes and phases w. r. t. ω and A are drawn in Fig. 9 and Fig. 10, respectively; using FOPD with $T_{0FOPD} = 0.622\ N*m$, the 3D/2D Bode plots are drawn in Fig. 11 and Fig. 12, respectively. From Fig. 9 and Fig. 11, and their enlarged 2D Bode plots Fig. 13(a) and Fig. 13(b), it is obvious that, the amplitude with the designed FOPD is bigger than that with the optimized IOPI, in low speed position tracking with the limit cycle as shown in Fig. 4(a) and Fig. 4(b). So, the tracking performance with the designed FOPD will be better than that with the optimized IOPI. At the same time, comparing Fig. 10 and Fig. 12, the phase delay with the designed FOPD is much smaller than that with the optimized IOPI, thus, the tracking system with the designed FOPD controller is more stable.

V. Experiment

In this section, extended experimental tests for the varying low speed position tracking are presented to validate the theoretical analysis. This experiment is performed on the same experimental platform – dynamometer, as in [2].

The varying low speed ($\pm 0.05\ rad/s$) position reference for tracking is shown in Fig. 14, and it can seen that, the position tracking performance with the designed FOPD in Fig. 15(b) is much better than that with the optimized IOPI in Fig. 15(a). It is more clear to see the difference of the position tracking performance in Fig 16(a) and Fig 16(b) for the position tracking errors, which can also be supported by the speed outputs in Fig 17(a) and Fig 17(b) for the IOPI and FOPD, respectively.

VI. Conclusion

In [2], simulation and experimental results show that the dynamic performance and robustness of the servo system tracking the normal speed with a designed FOPD controller outperforms that with the ITAE optimized IOPI controller. In this paper, using the describing function method and the Bode plots analysis, the observed advantage of the designed FOPD controller in [2] is explained for the nonlinear ultra low-speed position tracking system with friction effect over the optimized IOPI controller, which is consistently demonstrated by our extended experimental results.

Acknowledgements

Ying Luo would like to thank to the China Scholarship Council (CSC) for the financial support from Sept. 2007 to Feb. 2009 for visiting the Center for Self-Organizing and Intelligent

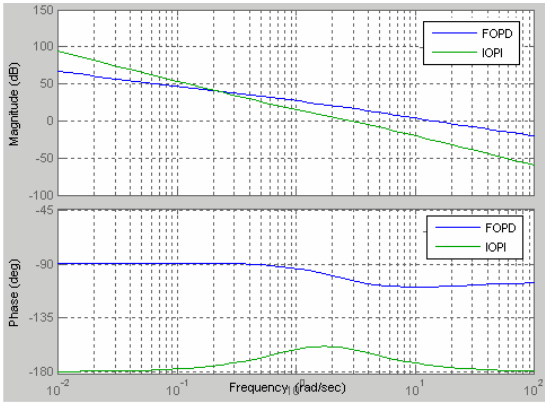


Fig. 8. Open-loop Bode plots without considering friction using the optimized IOPI and designed FOPD

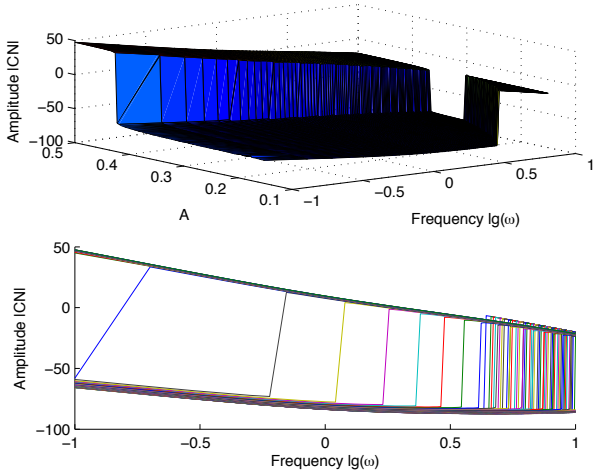


Fig. 9. 3D/2D Bode plot of the amplitude w. r. t. ω and A using IOPI with $T_{0IOPI} = 0.577\ N * m$

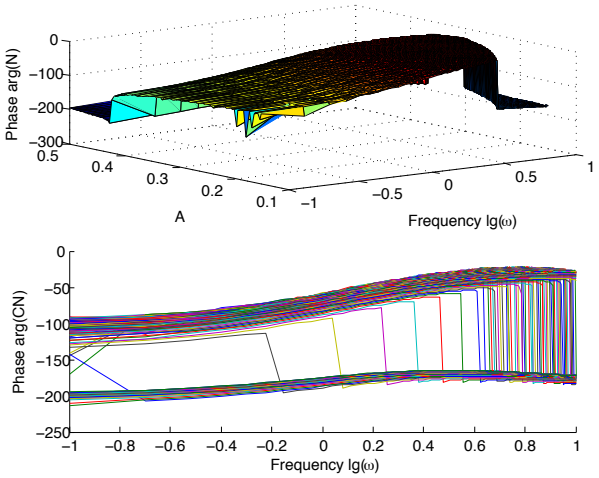


Fig. 10. 3D/2D Bode plot of the phase w. r. t. ω and A using IOPI with $T_{0IOPI} = 0.577\ N * m$

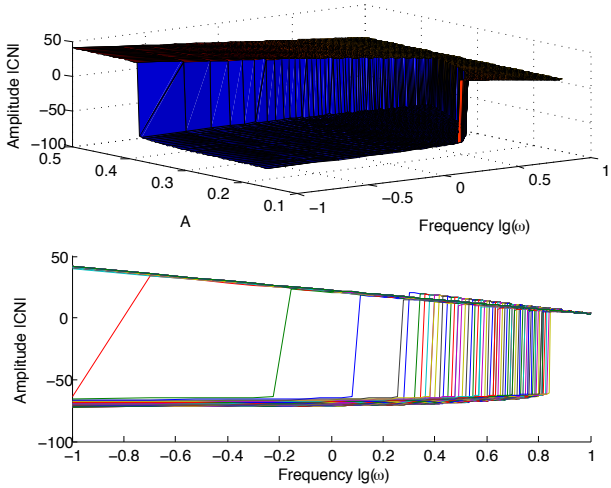


Fig. 11. 3D/2D Bode plot of the amplitude w. r. t. ω and A using FOPD with $T_{0FOPD} = 0.622\ N * m$

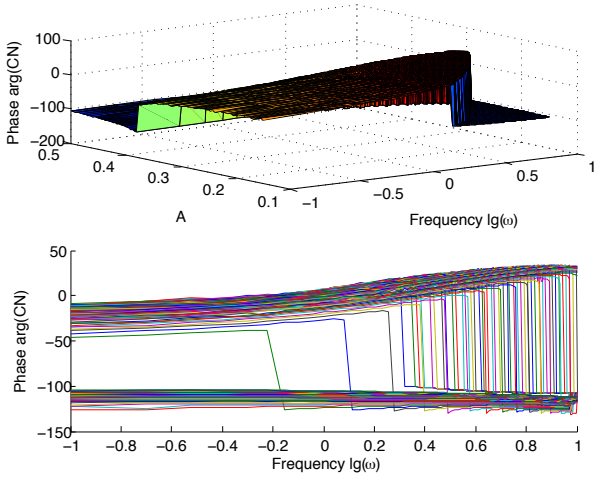


Fig. 12. 3D/2D Bode plot of the phase w. r. t. ω and A using FOPD with $T_{0FOPD} = 0.622\ N * m$

Systems, Dept. of Electrical and Computer Engineering, Utah State University.

REFERENCES

[1] H. Li and Y. Chen, "A fractional order proportional and derivative (FOPD) controller tuning algorithm," in *Proc. of Chinese Control and Decision Conference*, 2-4 July 2008, pp. 4059–4063.

[2] H. Li, Y. Luo, and Y. Chen, "A fractional order proportional and derivative (FOPD) motion controller: Tuning rule and experiments," *IEEE Transactions on Control System Technology (to appear)*.

[3] P. Lischinsky, C. C. deWit, and G. Morel, "Friction compensation for an industrial hydraulic robot," *IEEE Control Systems Magazine*, vol. 19, pp. 25–30, 1999.

[4] B. Armstrong-Hilouvry, P. Dupont, and C. C. de Wit., "A survey of models, analysis tools and compensation methods for the control of machines with friction," *Automatica*, vol. 30, no. 7, pp. 1083–1138, 1994.

[5] C. C. de Wit, H. Olsson, and K. J. Astraem, "A new model for control of systems with friction," *IEEE Trans. on Automatic Control*, vol. 40, no. 3, pp. 419–425, 1995.

[6] H. Olsson, "Describing function analysis of a system with friction," in *Proceedings of the 4th IEEE Conference on Control Applications*, 28-29 Sep. 1995, pp. 310–315.

[7] A. Mees and A. Bergen, "Describing functions revisited," *IEEE Trans. on Automatic Control*, vol. 20, no. 4, pp. 473–478, 1975.

[8] J. Tou and P. M. Schultheiss, "Static and sliding friction in feedback systems," *Journal of applied physics*, vol. 24, no. 9, pp. 1210–17, 1953.

[9] A. Nassirharand and S. R. M. Firdeh, "Design of nonlinear lead and/or lag compensators," *International Journal of Control, Automation, and Systems*, vol. 6, no. 3, pp. 394–400, 2008.

[10] F. B. M. Duarte and J. A. T. Machado, "Intelligent engineering systems and computational cybernetics," *Springer Netherlands*, 2009.

[11] M. S. Tavazoei and M. Haeri, "Describing function based methods for predicting chaos in a class of fractional order differential equations," *Nonlinear Dynamics*, 2008.

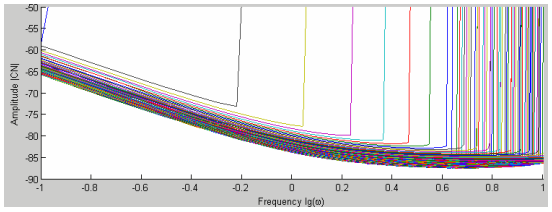
[12] K. S. Miller and B. Ross, "An introduction to the fractional calculus and fractional differential equations." Wiley, New York, 1993.

[13] I. Podlubny, "Fractional-order systems and $PI^{\lambda}D^{\mu}$ controller," *IEEE Trans. Automatic control*, vol. 44, no. 1, pp. 208–214, 1999.

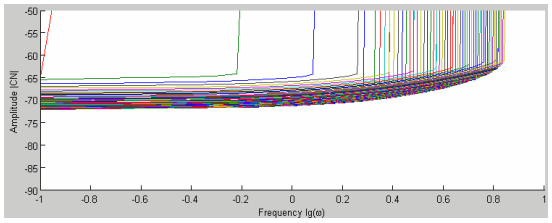
[14] D. Xue and Y. Q. Chen, "A comparative introduction of four fractional order controllers," in *Proc. of 4th IEEE World Congress on Intelligent Controllers and Automation (WCICA02)*, Shanghai, China, 2002, pp. 3228–3235.

[15] R. C.Dorf and R. H.Bishop, "Modern control systems," *Pearson Prentice Hall, Pearson Education, Upper Saddle River*, pp. 270–278, 2005.

[16] B. Armstrong-Helouvry, "Control of machines with friction," *Kluwer*



(a) Enlarged 2D Bode plot of the amplitude w. r. t. ω and A using IOPI with $T_{0IOPI} = 0.577\ N * m$



(b) Enlarged 2D Bode plot of the amplitude w. r. t. ω and A using FOPD with $T_{0FOPD} = 0.622\ N * m$

Fig. 13. Enlarged 2D Bode plot of the amplitude w. r. t. ω and A

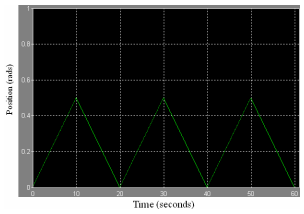
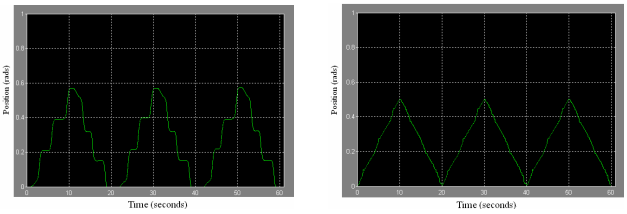


Fig. 14. Varying low speed position reference for tracking

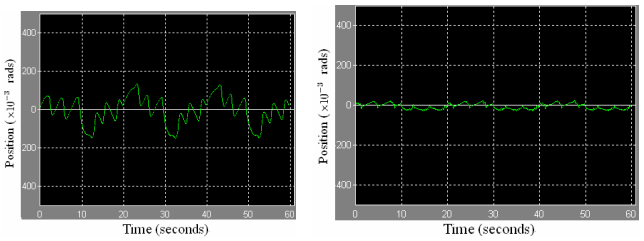
Academic Publishers, Norwell, MA, 1991.

[17] P. Dahl, "A solid friction model," *Technical Report TOR-0158(3107-18)-1, The Aerospace Corporation, El Segundo, CA*, 1968.



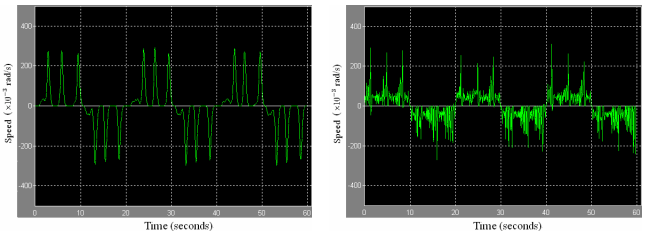
(a) IOPI (b) FOPD

Fig. 15. Position tracking outputs with varying low speed reference using IOPI / FOPD



(a) IOPI (b) FOPD

Fig. 16. Position tracking errors with varying low speed reference using IOPI / FOPD



(a) IOPI (b) FOPD

Fig. 17. Speed output with varying low speed position tracking reference using IOPI / FOPD